

ACCESS VERSUS CHOICE: TESTING THE 'FOOD DESERT' CONSTRUCT IN CHAMPAIGN, IL

BY

RHIANNON LEIGH JERCH

THESIS

Submitted in partial fulfillment of the requirements  
for the degree of Master of Science in Agricultural and Applied Economics  
in the Graduate College of the  
University of Illinois at Urbana-Champaign, 2013

Urbana, Illinois

Master's Committee:

Assistant Professor Kathy Baylis, Chair  
Assistant Professor Sahan T. M. Dissanayake, Colby College  
Assistant Professor Brenna D. Ellison  
Professor Craig Gundersen

## **Abstract**

How does one's food environment affect food purchase decisions? Food access has received significant political and academic attention, particularly under the Obama administration. Existing literature on food access and "food deserts" mainly focuses on geographic distance or the retail of fresh fruits & vegetables versus fast food within a neighborhood to determine and identify inequitable access. In this paper I develop an endogenous measure of food access by asking how geographic placement of food retail affects food expenditure, particularly of fruits & vegetables. I use novel data on 991 households matched to food prices from a census of geocoded food retailers in Champaign and Piatt counties to approach this question from two perspectives. I first estimate the household's share of grocery expenditures allocated to fresh, frozen, and canned fruits & vegetables versus other grocery items. I then use data on place of residence and geocoded data on food retail locations in Champaign County to test for relationships between retailer proximity, and the share of expenditure on fruits & vegetables. The next perspective uses a choice experiment to measure the tradeoff among store characteristics that determine where a consumer shops. The demand estimation reveals how much fruits & vegetables a person is actually consuming, while the choice experiment reveals whether that individual is constrained in their consumption by their existing food environment. I find that while proximity to a grocery store is positively correlated with healthier food consumption, policy response should focus on improving store quality and product quality to induce behavioral change. I further find policy response should be cognizant of endogenous locational sorting which may require alternative means to improve health other than changing the food geography.

## **Acknowledgements**

I wish to thank several people for their contributions in the development, analysis, and critique of this thesis: Daniel Karney, Taro Mieno, Catalina Londoño, Kashi Kafle and Olesya Savchenko for their generous time assisting me in understanding consumer theory and statistical programming; Bonnie Kelnberger for her crucial support and assistance in the world of mapping software; Professors Sahan Dissanyake, Craig Gundersen, and Brenna Ellison for their invaluable ideas, opinions, and criticisms; and Alexandra Hill, Jessica Pasciak, and Richael Young for their moral support and friendship. I especially wish to thank my advisor Kathy Baylis for her enthusiasm for research, unique ideas, and generosity. I am immensely grateful to have studied with her for two fun and productive years. I finally wish to thank my parents, Richard and Janie Jerch for their unconditional love and encouragement, and my best friend and love for eight years, James Pikul, who inspires me and challenges me every day. Thank you.

## Table of Contents

Chapter One: Introduction	
1.1 Motivation & Research Question .....	1
1.2 Hypothesis.....	3
Chapter Two: Review of Theoretical and Empirical Research	
2.1 Introduction.....	6
2.2 Non-Market Valuation.....	6
2.3 Literature on Food Demand.....	11
2.4 Contributions.....	18
Chapter Three: Methodology & Theoretical Framework	
3.1 Introduction.....	19
3.2 Demand System: The “Almost Ideal Demand System” .....	20
3.3 Choice Experiment.....	24
3.4 Robustness Tests.....	32
Chapter Four: The Data	
4.1 Data Collection.....	33
4.2 Summary Statistics: Demographics, Food Environment, and Budget Shares.....	33
4.3 Comparison of Preferences and Budget Shares.....	39
Chapter Five: Empirical Analysis and Results	
5.1 Introduction.....	42
5.2 Demand Estimation Results.....	42
5.3 Choice Experiment Results.....	49
5.4 Results of Robustness Tests.....	65
Chapter Six: Summary of Findings and Conclusions.....	
References.....	74
List of Figures.....	79
Appendix A: Choice experiment design SAS macro and SAS fractional factorial design output.....	89
Appendix B: Conditional Logit regression results.....	90
Appendix C: List of produce for which respondents provided consumption information.....	91
Appendix D: Distribution of choice experiment selections based on the produce quality attribute.....	92
Appendix E: MMNL results with Main Effects – Only specification.....	93
Appendix F: <i>Weekly Grocery Bill</i> estimators linear combination significance test by respondent group.....	96
Appendix G: Regression results of Robustness Tests.....	97

## Chapter One: Introduction

### 1.1 Motivation & Research Question

Outside of a Save-a-Lot store in Chicago's Auburn Gresham neighborhood in Spring of 2011, mayor Rahm Emanuel praised a joint initiative between the city and the retailer to stamp out food deserts: "The notion that people do not have access to fresh fruits and vegetables is unconscionable, and you cannot have that today" (*Time Out Chicago*). The United States Department of Agriculture (USDA) defines "food desert" as a low-income census tract where a significant share of residents live more than one mile from a large grocery store. Roughly one in five Chicagoans lives in a food desert, according to a 2006 study by Mari Gallagher Research & Consulting Group.<sup>1</sup> One such area includes Auburn Gresham, a heavily African-American neighborhood on Chicago's south side. After Emanuel's speech, resident Taylor McDowell stated in an interview with *Time Out Chicago*: "...[T]his is not a food desert: We have a Jewel down the road, another Save-A-Lot on 79<sup>th</sup> and Halsted, Aldi around the corner on State. People shouldn't feel sorry for us. This is a black middle-class neighborhood. It's not like I'm suffering."

Today, federal, state and municipal governments are rushing to vanquish "food deserts" across the country. In February of 2010, the Obama Administration proposed a \$400 million Healthy Food Financing Initiative to promote retailers of healthful foods to move to underserved urban and rural communities. Since 2001, California, Nevada, New Mexico, Texas, Oklahoma, Louisiana, Illinois, New York, Pennsylvania, D.C., and Maryland have enacted legislation aimed at promoting healthier food retail. Such legislation has included financial incentives to attract healthier food retail outlets to underserved areas, improving healthy food offerings in existing stores through subsidized remodeling and restocking, and relaxing zoning requirements to make it easier for grocery stores to locate in underserved areas (Centers for Disease Control and Prevention). Such policy interventions rest on the assumption that the mile-or-more expanse to a grocery store prohibits area residents from consuming produce and healthy foods. This paper creates a framework to investigate whether residents of these so-called "food deserts" are truly constrained by their food geography.

Classical consumer theory purports that individuals are mobile and possess perfect information about choice alternatives, including residential location. As the Tiebout model suggests, people move and are free to choose their communities (Tiebout 1956). Still, there is evidence that a dearth of grocery stores is more likely to occur in poorer areas (Berg and Murdoch 2007; Ver Ploeg et al. 2009; Morland et al. 2002, Liese et al. 2007). Residents of poorer neighborhoods may face challenges to food access on several fronts: convenience stores and quick-service restaurants tend to outnumber grocery stores (Larson, Story, & Nelson 2009; Morland et al. 2002); fresher nutrient-rich foods tend to be more expensive than calorie-dense, nutrient-poor foods; (Newby 2006; Drewnowski 2003; Monsivais & Drewnowski 2007), and those retailers that provide lower-cost health options fail to establish

---

<sup>1</sup> 500,000 people cited in the study divided by the 2000 US Census Chicago population of 2.9 million.

stores in low-income areas due to security costs or high profit risk. While prohibitive supply-side costs can deter grocery stores from locating in low-income areas, so can demand-side factors such as consumer preferences. A food desert might arise in a geographic area with insufficient supply of nutritious food, yet a food desert might also arise in an area with insufficient demand (Bitler & Halder 2011).

In either case, the extra distance required to purchase fresher foods increases the implicit costs of these foods. The result is that some people face greater chances of obesity and diet-related disease while others may regularly shop for their food at convenience stores. Recent research has found a positive correlation between supermarket access and diet and health (Wilde, Gundersen, Baylis, & Llobrera 2012; Larson et al. 2009; Chen, Florax & Snyder 2010) as well as a negative correlation between convenience store access or fast food access, and diet and health (Wilde et al. 2012; Giskes, van Lenthe, Avendano-Pabon, & Brug 2011; Gibson 2011; Currie, Della Vigna, Moretti & Pathania 2009). However, the policy response to “food deserts” has largely ignored the fact that a causal relationship between access and diet has yet to be proven. Researchers in the United Kingdom, for example, found that cultural influences represented by a person’s age or gender are significant determinants of healthy food consumption whereas distance to a grocery store, produce prices, or socio-economic status are not (Pearson, Russell, Campbell, and Barker 2005). If policy response aims to promote dietary health, the appropriate action depends on solving an identification problem: whether poor diet and obesity result from supply-side factors, or from an individual’s inherent preferences.

The food policy and public health literature has called significant attention to the proper definition and identification of a “food desert” (California Center for Public Health Advocacy 2007; Mari Gallagher 2006; Ver Ploeg et al. 2009). The USDA *Food Access Research Atlas* identifies over 8,800 census tracts, harboring 29.7 million people, which qualify as food deserts.<sup>2</sup> By incentivizing grocery stores to locate in these areas, policy makers, including the Obama administration and Rahm Emanuel, have reacted under the presumption that insufficient access to grocery stores deters people from eating fresh produce. Further, they expect access will have an opposing effect, and induce residents living nearby a grocery store to consume more fresh produce.

There are two crucial limitations to this approach. First, it assumes food geography is exogenous to the household, and differences in consumption behavior can be attributed to observable differences in proximity to a grocery store. In other words, the “food desert” construct ignores the Tiebout principle that people may choose where they live. Consumption behavior may be exogenously driven, but may also be a product of pre-existing consumer predilections and tendencies. In particular, endogenous locational sorting may drive those who enjoy fresh produce to live nearby a grocery store, such that the apparent positive effects of grocery store proximity on diet are spuriously related. Recent research on locational sorting provides evidence that people can, and sometimes

---

<sup>2</sup> <<http://www.ers.usda.gov/data-products/food-access-research-atlas/documentation.aspx#.UUvLCBk1cYJ>>

do, choose their housing location according to their preferences for amenities and environmental attributes at those locations (Bayer and Timmins 2005). The second major limitation of current policy response is the primary focus on *availability* of fresh produce and produce retailers, leaving *quality* of produce and retailers as a secondary concern at best, and non-concern at worst. Research informing food access policy draws mainly on qualitative data such as location and size and ignores quality or variety of food entirely (Bernstein 2012). This paper explores potential weaknesses of the “food desert” construct by asking first whether food geography can constrain consumption of fresh produce; and second, whether the source of such consumption constraints may be exogenously or endogenously driven.

In 2008 and 2009, Detroit received significant national attention from academics, the media, and the federal government for its dearth of grocery stores. Detroit resident Noah Stephens initiated a photography survey project to document the produce of Detroit’s supermarkets. Stating his motivation for the project, Stephens conceded: “The food choices of many Detroiters may be constrained by affordability, transportation, safety, food quality. Or people may just like potato chips. The heart of this project is the question of access versus choice” (Huffington Post 2012). Like Stephens’ *Kickstarter* project, this paper seeks to address and quantify access versus choice.

## **1.2 Hypothesis**

In this paper I attempt to develop an endogenous measure of food access by asking how consumer preferences interact with geographic location of food retailers to affect food consumption, particularly of fruits & vegetables. I use novel data on 991 households matched to food prices from a census of food retailers in Champaign County to approach this question from two perspectives. I first estimate household shares of grocery expenditures allocated to fresh, frozen, and canned fruits & vegetables and other grocery items. I then use data on household location and geocoded data on food retail locations in Champaign & Piatt counties in central Illinois to test for correlation between retailer proximity and household grocery expenditure share of fruits & vegetables. The second perspective uses a choice experiment to measure preferences for grocery store characteristics that determine where a consumer shops for food. I then compare the grocery expenditures and preferences between individuals who live near a grocery store, and those that live further from a grocery store. If these two subsets have differing consumption behavior, but similar preferences, this will suggest food geography constrains consumption. Such a result would support current policy’s assumption that exogenous “treatments” of adding grocery stores to “food deserts” has the causal effect of inducing fresh produce consumption. On the other hand, if both consumption behaviors as well as preferences are systematically different between individuals living nearby a grocery store and individuals living far from a grocery store, this will suggest food geography has an indeterminate effect on food choice, and instead endogenous sorting may be the primary cause of divergent consumption behavior. Alternative policy responses such as subsidizing fresh produce or nutrition education may be more effective at combating poor food choice than simply adding a grocery store.

Fruit & vegetable consumption is the main focus of this thesis for two reasons. First, nutritional science generally agrees that eating fruits & vegetables is a primary component of dietary health and is associated with reduced chronic diseases (U.S. Department of Health & Human Services). Second, consumers who are food-access constrained are more likely to compromise consumption of fresh foods compared to non-perishable foods because either the fresh foods are expensive relative to non-perishable foods, unavailable entirely, or poor quality (Drewnowski, 2003; Drewnowski, Darmon, & Briand 2004; Monsivais & Drewnowski 2007; Bitler and Haider 2011). Fresh fruits & vegetables are costly to supply because stocking fresh produce necessitates a refrigerated and timely supply chain as well as expensive in-store refrigeration, which both serve to increase the sale price of fresh produce. This is particularly true for if the retailer supplies a small quantity and is unable to harness cost savings through economies of scale (Bitler and Haider 2011; Ver Ploeg et al. 2009). The effective cost of fresh produce, thus, increases in absolute terms, but also in terms of opportunity cost. A consumer who must travel far to grocery shop will want to maximize their grocery expenditure and purchase foods that will last until their next grocery shopping trip. Thus perishable foods are effectively more expensive to an infrequent grocery shopper because the shelf life opportunity cost of the fresh food is high relative to the non-perishable foods. The shelf life opportunity cost issue exacerbates with poor quality produce, which may be both displeasing to the consumer, and can spoil quickly. Geography can therefore constrain a consumer's effective budget, time availability, and quality preferences. Consequently, consumers will be more likely to trade off fresh fruits & vegetables for non-fresh fruits & vegetables in order to reduce costs in a restrictive food environment.

I find that survey respondents affiliated with the University of Illinois exhibit more inelastic demand for fresh produce the closer they live to a grocery store. Additionally, respondents who live within one mile of a grocery store have lower propensity to substitute fresh for non-fresh produce compared to respondents who live over one mile from a grocery store. While endogenous sorting explains some of this divergent consumption behavior, a significant proportion of respondents who live over one mile from a grocery store would likely exhibit more stable consumption of fresh produce if higher quality produce and specialty stores were made accessible to them.

This study extends the current literature on food choice on multiple fronts. First, this thesis contributes specifically to existing empirical demand work by utilizing first-hand data on food purchases and observed prices to estimate fresh produce demand in Champaign & Piatt counties, Illinois. Further, this study contributes to food access policy by empirically testing for consumption constraints on multiple fronts, including product quality, and retailer type, in addition to distance and cost. Current food access policy relies on geographic measures, such as number of large grocery stores per square mile, to identify "food deserts", or areas deemed inhospitable to healthful food consumption. This paper allows individual preferences, as opposed to hetero-normative definitions of acceptable food landscapes, to motivate identification of areas truly underserved. This research can improve targeting areas



in genuine need of policy intervention through quantifying gaps in consumers' preferences and their actual consumption bundles.

In the next chapter, I review theoretical and empirical literature concerning food access, hedonic analyses, and food demand estimation. In the third chapter, I present a theoretical basis for evaluating consumer preferences and consumer behavior through a choice experiment and demand system. In the fourth chapter, I discuss my data and provide summary statistics. The fifth chapter discusses my results and conclusions, and the final chapter summarizes and concludes my findings.

## **Chapter Two: Review of Theoretical and Empirical Research**

### **2.1 Introduction**

This thesis investigates how the food environment affects food choices in Champaign & Piatt counties, IL. Specifically, I examine how the array of foods available to you affects, if at all, what you purchase. I employ both demand system estimation as well as a choice experiment to answer this question with particular concern for fruit & vegetable consumption. Analysis of microeconomic food demand and consumption behavior at the disaggregated level possesses an inherent simultaneity problem; demand is affected simultaneously by the choice of where to shop - a discrete choice - and what to consume - a continuous choice. When purchasing groceries, the consumer cannot make one decision without implicitly making the other decision. Previous literature has looked at both pieces in isolation, but few scholars have considered the discrete choice and the continuous choice as simultaneous determinants of food consumption. In order to understand how, and to what extent, the food environment affects healthy food consumption, I investigate both the choice of where to grocery shop, as well as the choice of what foods to consume once at the grocery store. I compare these choices between survey respondents living within one mile of a grocery store, and respondents living over one mile from a grocery store based on the USDA geographic criteria for residence of a “food desert.” I will hereinafter refer to these respondent groups as “1 mile or less” residents and “Over 1 mile” residents, respectively.

Two questions drive the methodology of this paper:

- 1) How are the preferences that drive where people grocery shop - the discrete choice - different among “1 mile or less” residents and “Over 1 mile” residents?
- 2) How does consumption of fresh fruits & vegetables, non-fresh fruits & vegetables, and other foods – the continuous choice - differ among “1 mile or less” residents and “Over 1 mile” residents?

In the following sections, I review literature concerning theoretical and empirical applications of choice preferences and consumption behavior that will provide a basis for my subsequent analyses.

### **2.2 Non-market Valuation**

#### **2.2.1 Introduction**

In order to understand preferences that drive where people grocery shop, I consider what people are willing to pay for their preferred grocery shopping experience. Specifically, I consider how individuals might value improved access to groceries. Food “access” possesses several dimensions in addition to proximity. A corner store next door to your house might carry apples, but if these apples are bruised, tasteless, and expensive, you will be unlikely to purchase them. It follows that I can consider “improved access” to include product appearance, quality, and price, in addition to proximity. Further, assume two supermarkets selling apples of the same quality and price exist next door to one another. If Store A has substantially better customer service, shorter check-out lines, or sources its produce in a more sustainable manner than its competitor, Store B, you may be more inclined to shop at Store A.

Harold Hotelling argued that even if Store B cut its apple prices, many consumers will still prefer Store A because preferences for in-store service, merchandise mix, product quality, methods of doing business, and political persuasions of the seller or vendor can all shield Store B from harsh price competition with Store A (Brown 1989). An apple purchased at any grocery store has several characteristics that dictate whether or not you purchase that apple, besides those specific to the apple itself. Even for completely standardized products, absolute homogeneity does not exist if we account for the location and characteristics of the retailer. Thus, in deciding what groceries to purchase, we inherently consider not only store location and the time required to travel to the store, but also product quality, product price, and the marketing chain in which the groceries are sold, among other things (Hanemann 2006). The result is that an apple purchased at one store location is not the same as an identical apple purchased at a different store because the time it took to acquire the apple, its price, the store ambience, and the marketing channel from which it came can all differ. In this thesis, I investigate if and how people alter their produce purchases given variations in these characteristics.

### **2.2.2 Lancaster's "New Theory"**

Lancaster postulated that utility, or a preference for one good over another, is derived from characteristics that goods possess rather than the goods themselves (Lancaster 1966). Kelvin Lancaster's "new theory" provides the theoretical basis for estimating changes in utility that result from attribute changes of a good. His theory holds that the good's characteristics are taken as given, and the consumer is free to vary only the quantity they consume of a particular good. Lancaster postulated that the more advanced an economy, the more characteristics available relative to goods, so that in the most advanced economies, consumers could obtain any combination of characteristics they desired, constrained only by their budget.

Product characteristics cannot be purchased individually by the consumer. For example, a consumer cannot buy high quality produce separately from a five-minute travel time and manufacture their own grocery store. Grocery stores comprise "lumpy" goods because shoppers have a finite, discrete opportunity set of grocery stores to patronize, or "consume." In other words, the array of grocery store characteristics available to a consumer will not be an exhaustive set of possible store attribute combinations. Assuming that travelling to multiple stores is costly, consumers must decide which of these characteristics is most important, and patronize the store that best fits their criteria. By isolating the consumer's internal tradeoffs of grocery store characteristics, such as travel time, produce quality, price, and marketing channel, it is possible to assign a numeric value or "price" to these characteristics through non-market valuation. Non-market valuation measures the satisfaction from these store characteristics as the monetary amount that a consumer would be just willing to exchange for the characteristic if it were possible to make such an exchange (Hanemann 2006). Non-market valuation enables us to understand how consumers value different store characteristics, and in turn, predict the store people are most likely to patronize

assuming any given store type is accessible to the consumer. This paper aims to identify how quantity of fruit & vegetable consumption is affected by varied access to an individual's preferred grocery store.

### **2.2.3 Contingent Valuation & Choice Experiments**

Preferences for particular good characteristics can readily be observed through the monetary value individuals place on the consumption or use of a good or service. Contingent valuation is a form of non-market valuation that utilizes interviews or surveys to elicit individuals' monetary value for particular goods (Hanneman 2006). If products are considered to be composed of attributes, such as the Lancaster theory suggests, a form of contingent valuation known as a choice experiment can be used to determine which product attributes are important to overall preferences for that product and which combinations of attribute levels are most preferred. Using rankings or rating-scale evaluation judgments of products, choice experiments decompose such rankings into components based on qualitative attributes of the products (Kuhfeld 2010).

Choice experiments are a popular contingent valuation technique in marketing, environmental, and public policy research (Kuhfeld 2010; Louviere, Hensher & Swait 2000; Alpizar, Carlsson, & Martinsson 2001). A choice experiment provides a framework in which one can establish a set of behavioral rules under which it is reasonable to represent an individual's choice-making decision when considering a set of alternatives. Choice experiments estimate welfare changes resulting from changes in a good's attributes. In the case of this paper, the choice experiment estimates changes in consumer welfare for changes in grocery store attributes, where "attributes" are defined as the characteristics of a store that cause a person to choose one alternative over another. Attributes are the source of utility (Hensher, Rose, and Green 2005).

My thesis applies Lancaster's concept of characteristic-derived utility to the decision of where to grocery shop for residents of Champaign & Piatt counties. I hypothesize that unless consumers can obtain their preferred combination of grocery store characteristics, they are constrained in their ability to purchase the exact bundle of food products they desire because the decision of what groceries to purchase is wrapped up in a simultaneous decision of where to grocery shop. I test whether consumers are constrained in their consumption of produce through a choice experiment framework where grocery stores are considered to be composed of attributes including size, price, service offerings, produce quality, and travel time. For example, one hypothetical grocery store may possess large floor space, a pharmacy and florist, and a wide variety of non-food items, sell high quality fresh produce, require five minutes of travel time to access, and charge a five percent premium for groceries. While, another hypothetical grocery store may possess small floor space, sell very few non-food items, offer a limited array of fresh produce, require 15 minutes of travel time to access, and charge a 15 percent discount for groceries. The choice experiment is used to determine which attributes are important to food product preference, and which combinations of attributes are most preferred among respondents (Kuhfeld 2010).

In reality, most consumers face multiple constraints that restrict them from choosing their preferred alternative. An individual may not be able to afford a car that they want because of expense. Similarly, residents of Champaign County, though they prefer to grocery shop at Trader Joe's, are inhibited by the 2.5-hour travel time needed to get to the nearest store.<sup>3</sup> Yet, through the choice experiment I can allow such constraints to be non-binding in order to elicit information about a person's preferences.

#### **2.2.4 Grocery Store Preferences: Previous Empirical Work**

Literature on consumer preferences for supermarkets have explored the effects of store size, travel distance, customer service, and quality on grocery shopping behavior. Food access policy has historically used quantifiable metrics such as store size or travel distance to identify areas where residents cannot easily consume fresh produce. However, researchers Blitstein, Snider, and Evans (2011) found that grocery shoppers don't only care about cost and proximity, but also need choice and quality if they are going to buy fresh produce. Their study utilized survey data from 495 respondents in a low-income neighborhood of Chicago, IL to examine how residents' perceptions of their food shopping environment was related to their dietary intake of fresh produce. The researchers found that among their urban, minority low-income sample, those who felt their produce vendors provided high quality produce and selection consumed higher quantities of produce. In a scientific health study of retail food environments in northern Illinois, Izumi et al. (2012) noted that lack of safety or cleanliness and poor customer service could deter shoppers from consuming healthful foods.

Daniel Brown (1978) found in a panel study of 101 Chicago suburban women that travel distance to a grocery store had a stronger effect than store size in determining store patronage. Brown also found that grocery store size attracts patronage up to a point, after which increasing size deters grocery shopping. People want variety, but not too much. Gómez, McLaughlin, and Wittink (2004) found that a decrease in product quality (as opposed to an increase), and changes in customer service and value explain 75 percent of the variation in customer satisfaction among consumers surveyed in 250 east coast stores. Thus, product quality and customer service are important determinants of utility with respect to grocery shopping.

Huang and Oppewal (2006) conducted a choice experiment among 152 supermarket shoppers in England to determine which attributes are most instrumental in causing people to purchase their groceries online rather than in-store. Retailer attributes included the online shopping delivery charge, time availability of the respondent, travel time to the grocery store, and purpose of the grocery purchases. Travel time has a stronger effect on the decision to purchase groceries online than does the delivery charge, which indicates people place high value on their personal time. The authors further find that the purpose of the shopping trip has little effect on respondent's

---

<sup>3</sup> Of 886 complete responses, 60 noted they would shop at Trader Joe's if one existed in Champaign County.

choice outcome. People will shop for ingredients for a dinner party at the same place they will shop for their routine groceries.

Analyses of grocery store preferences appear frequently in the marketing literature, as well as the urban planning literature. Arnold, Oum, and Tigert (1983) investigated the “determinant attributes” of store choice, which they defined following Myers and Alpert (1968) as the “attributes projected by the product’s image which lead to the choice of that product (p. 150) [.]” Using a multinomial logit model, the authors identified robust determinants of grocery store patronage through interview-based survey data across six geographic markets and six years.

Locational convenience and low prices were the most important determinant attributes across geographic markets and over time. In addition to store location, Tang, Bell and Ho (2001) also found that familiarity with the store, service quality, and breadth and depth of products are key drivers of shopping utility. The authors parsed a utility model of fixed and variable costs and benefits to describe overall grocery shopping utility. Using data from 500 households across five stores over two years, their utility model correctly predicted 83 percent of the sample’s store choice. Carrasco (2008) perceived determinants of grocery store choice as a destination choice problem in transportation planning. He used a revealed preference approach and travel behavior survey data on 32,000 Swiss households from 2005 to investigate what factors best predict where individuals choose to grocery shop. The author’s discrete choice model included variables for the following store characteristics: store quality, price range, store size, opening hours, distance from home, distance needed to deviate from route for non-home based trips, and retail density near the grocery store. He found households have a clear preference for larger stores. He also found sample households are averse to travelling far from home to grocery shop, and greatly dislike deviating from a set route in order to purchase groceries. Higher income households had a stronger aversion to travel distance, indicative of the value these households assign to time. Also, households preferred relatively low retail density near their grocery store, likely due to time savings from less congestion on the shopping errand.

A 2011 revealed preference study by Hsieh and Stiegert (2011) found that households that spend a high proportion of their grocery bill on non-perishable grocery items are more likely to shop at value-oriented stores such as supercenters, and are less likely to shop at high-end, natural food supermarket chains. Their study gives evidence that travel distance from a consumer’s household to the store is an important determinant of the store type in which a consumer decides to shop. My thesis research aims to take one step further to investigate whether distance also affects the foods consumers ultimately purchase.

## 2.3 Literature on Food Demand

### 2.3.1 Introduction & Theoretical Considerations

“Food deserts” are problematic to public health in so far as they inhibit consumption of fresh produce. If food geography inhibits or induces consumption of fresh produce, its effect will manifest through distinctive consumption behavior between “1 mile or less” residents and “Over 1 mile” residents.

The basic objective of consumer behavior theory is to explain how a rational individual makes decisions on what to consume when faced with a constraint. Under the assumption that a consumer can make a complete and consistent order of preferences for particular goods within a finite choice set, the choices the consumer will make are limited by their income or budget (Bentham 1789). The consumer must find an optimal bundle of goods within the finite choice set so as to maximize their utility but keep expenditures at or below their budget constraint (Teklu and Johnson 1986). A budget constraint is one example of factors that shape, or “constrain” the possibility frontier within which preferences can be honored. One may equally think of additional constraints such as time and location that limit the combinations of goods that an individual can consume. For example, current policy addressing food deserts asserts that residents face both budgetary and geographic constraints in their grocery consumption decisions (Ver Ploeg et al. 2009). Current policy is particularly concerned with how such constraints inhibit demand for fresh, healthy foods. Because demand for fruits & vegetables is integrated with demand for other types of foods if both are considered a subset of the same commodity group, I investigate relationships between food geography and fruit & vegetable demand using a system of demand. The system measures fruit & vegetable demand simultaneously with all other food demand.

Figure 1 illustrates individual consumer demand through a graphical representation. The y-axis is a ratio of quantity of two goods, and the x-axis is a price ratio of the two goods. The horizontal line is a budget constraint representing all possible combinations of the two goods that the consumer can purchase subject to a finite amount of money they can spend. If the consumer’s budget constraint changes, the amount of the two goods they consume may change. Further, if the price of one good changes, the ratio in which they consume both goods can also change. Thus, variation in an individual’s budget can change their consumption levels, and changes in prices of each good can change the ratio in which both goods are consumed. Otherwise stated, changes in prices can cause the consumer to “trade off” one good for the other.

Quantification of trade off behavior is measured through elasticity of income and the elasticity of substitution (or cross-price elasticity), respectively. Income elasticity measures changes in quantity purchased given a change in the consumer’s budget, where cross-price elasticity measures changes in consumption of a particular good given changes in price of a different good. Cross-price elasticities are generally negative for complementary goods and positive for substitute goods. Own-price elasticity, also referred to simply as the price elasticity of demand,

measures the percent change in consumption of a good given a percent change in price for the same good. Price elasticities are generally negative, except in the case of Giffen goods (Nicholson & Snyder 2008, p159). Price elasticities are necessary in order to understand consumer behavioral responses to variation in product price. Price elasticity of demand is determined by several factors including availability of substitutes, income, consumer preferences, and household budget share for a particular product, among others (Nicholson and Snyder 2008, p. 150).

Demand is inelastic when the relative change in quantity purchased is less than the relative change in price, so that the absolute value of the own price elasticity is between 0 and 1. Demand is elastic when the relative change in purchased quantity is greater than the relative change in price, so that the absolute value is greater than 1. Price elasticity of demand, itself, is dimensionless, but provides a reference for comparing purchasing behavior for price changes across different goods. Goods that are commonly price inelastic include necessity or every-day-use items with few substitutes such as public transportation, cigarettes, staple food items like eggs and rice, or gasoline. Price elastic goods have relatively many substitutes. Price elastic goods can include “luxury” goods, not necessary for every-day-use such as cinema tickets, imported French wine, or vacation travel. Because price elasticity lacks dimension and provides meaning only in relative terms, the degree of elasticity for a particular good depends on an individual’s theorized consumption basket (Nicholson and Snyder 2008 pg. 26).

I estimate both uncompensated, or “Marshallian”, and compensated, or “Hicksian” demand elasticities. Where compensated demand models estimate only substitution between products, uncompensated demand models account for both substitution effects and the effective income change resulting from a good’s price change (Nicholson & Snyder 2008, p. 151). I further use both income elasticity as well as own-price elasticity to compare behavioral differences in grocery consumption between “1 mile or less” and “Over 1 mile” residents. Income and price elasticities will reveal whether proximity to a grocery store is correlated with greater price sensitivity of fresh produce. Cross-price elasticities will reveal whether food geography is correlated with greater substitution of fresh for non-fresh produce.

### **2.3.2 Food Demand Functional Forms**

The Almost Ideal Demand System (AIDS) pervades the empirical food demand estimation literature. The model was first developed and applied to consumer demand empirics by Deaton and Muellbauer in 1980. The AIDS model, unlike simpler demand models based on the linear expenditure system generalization of Cobb-Douglas (e.g., Rotterdam model, translog models, etc.) does not require that consumers at all points along the income distribution allocate their budgets identically (Nelson 2013; Tiffin, Balcombe, Salois, and Kehlbacher 2011). The resulting system is flexible, thus better situated to model empirical data. Further, the resulting system satisfies the neoclassical utility theory axioms of choice exactly, including the adding up, homogeneity, and symmetry



restrictions (Deaton and Muellbauer 1980a). The model gives an arbitrary first-order Taylor series approximation to any demand system, and conditions of the model guarantee the existence of theoretically consistent aggregate demand (Deaton and Muellbauer 1980a; Nelson 2013), an advantage over the Rotterdam model perpetuated by Henri Theil and Anton Barten in the 1960's. The AIDS model allows the theoretical adding-up, homogeneity, and symmetry restrictions to be tested statistically rather than imposed, as with translog linear demand models (Deaton and Muellbauer 1980a). Additionally, estimation of the Rotterdam model requires time series variation in the data, either through panel data or times series on aggregate consumption (Nelson 2013). Because data for this thesis comprises a cross-section, use of the Rotterdam model would be inappropriate.

Determination of demand system matrix rank provides the precursor to appropriate model specification. Rank 1, rank 2, or rank 3 imply budget shares and total expenditure exhibit either an independent, linear, or quadratic relationship, respectively. A rank 2 system, for example, has budget shares that are linear in log total expenditure and are composed of two independent price functions. Gorman (1981) demonstrated that any exactly aggregable demand system could possess at most three independent income (or expenditure) terms in order to satisfy Slutsky symmetry,<sup>4</sup> such that the highest possible rank is rank 3.

The rank 2 AIDS form is generally consistent with empirical food expenditure data. For example, Katchova and Chern (2004) tested several demand models using Chinese household expenditure data on grains, vegetables, fruits, pork, and poultry and concluded the AIDS model produced the most reasonable set of expenditure elasticity estimates based on their a priori expectations compared to the rank 3 model. Wood, Nelson, and Nogueira (2009) estimated consumer welfare loss resulting from increased staple food prices among poor households in Mexico. Their Engle curve analysis based on food expenditures and budget shares across six food groups identified rank 2 as appropriate for estimating Mexican food demand among low-income households. Huang and Lin (2000) derived a modified version of the AIDS model to estimate food demand using 4,500 American households. The authors assumed linearity in income, thus did not test model fit using quadratic income terms. They investigated elasticity differences among three income brackets. The majority of their demand elasticities were statistically significant and met a priori expectations on size and magnitude. Tiffin et al. (2011) examined how changes in food prices and food expenditure affect consumption of various nutrients among UK households. Their extensive literature review of European household food demand estimation concluded that the AIDS model was most robust. Eales and Unnevehr (1994) found the AIDS model has a functional form consistent with observed US meat demand. Erdil (2006) compared estimates for agricultural products in OECD countries between the AIDS, CBS (Central Bureau of Statistics of the Netherlands), and Rotterdam models using a nested model selection procedure. For 12 of the 25

---

<sup>4</sup> The Slutsky equation (also known as the Slutsky identity) demonstrates that change in the demand for a good caused by its price change can be decomposed into two effects, the substitution effect and the income effect. Slutsky symmetry must be satisfied in order to relate changes in uncompensated demand with changes in compensated demand and empirically estimate demand elasticities (Nicolson and Snyder 2008, p. 156).

OECD countries, the AIDS model provided the best representation of demand for agricultural products. Lee, Brown, and Seale (1994) tested several models combining features of the Rotterdam and AIDS models to estimate Taiwanese consumer demand for seven share categories including food, and found the AIDS-based models better described consumer behavior.

Conventional demand system estimation techniques that specify budget shares as the dependent variable, including the AIDS model, have received recent criticism regarding use of unit values as a proxy for price (Gibson and Rozelle 2005; McKelvey 2011). When price data for particular goods is not available, which is often the case, price data is inferred using expenditure and quantity survey data to generate unit values. One concern is that individuals may respond to price changes by substituting on the quality margin rather than quantity, which would understate the change in unit value. For example, a consumer may respond to a price increase of fresh organic green beans by substituting the fresh beans for frozen, conventionally grown green beans. When the price of fresh beans increases, the change in unit value for the category containing green beans will understate the original price increase, all else equal. The imperfect proxy for price can be exacerbated if the food category is particularly broad, if several quality differentials are available to the surveyed consumer, or if the household sample contains significantly heterogeneous consumers. Wealthier households, for example, may exhibit high income- elasticity of demand for particular foods because of frequent substitution on the quality margin (Moro, Daniele, and Sckokai 2000; Abdulai 2002). McKelvey (2011) demonstrated that quality substitution is substantively large when comparing demand estimation results between unit value-based estimation and actual price data.

I use a sampling of actual price data for each food category to avoid any issues arising from use of unit values. Additionally, I disaggregate food categories using quality as a major characteristic, thus mitigating major quality differences stemming from aggregation. For example, I separate fresh produce from frozen and canned fruits or vegetables. Separation of these food categories follows Deaton (1988) whereby commodity groups are deemed to be weakly separable. Weak separability is a necessary condition for two-stage budgeting in which the consumer first decides how much to spend on food, clothing, and other broad product groups, then allocates consumption of particular goods within those broad groups. Consumption of goods within each group is modeled as a sub-utility function, and a sub-maximization problem. The weak separability assumption is important for demand estimation because it places severe restrictions on the degree of substitutability between goods in different groups. While a consumer may substitute between fresh green beans and frozen green beans, we assume they do not substitute between fresh green beans and shoes. Further, by assuming that commodity groups are weakly separable, considerable structure is imposed on the nature of quality responses to price changes (McKelvey 2011).

### 2.3.3 Review of Fruit & Vegetable Demand & Consumption Behavior

Consumer theory accepts that food demand in industrialized countries is largely inelastic, or price-insensitive (Herrmann and Roeder 1998). The USDA *Commodity and Food Elasticities* database reports average income elasticity for fruit & vegetables in the US of 0.22 and an average own price elasticity of (-.26), both relatively inelastic indicators.<sup>5</sup> You, Epperson, and, Huang (2011) estimated a demand system for 11 fruits and 10 vegetables using aggregate national data from 1960 through 1993 and found fresh fruits & vegetables respond significantly to changes in their own prices but less so to changes in income, suggesting that produce price has greater influence than household income in determining fruit & vegetable demand among US consumers. Income elasticities of demand for fresh vegetables & fresh fruits were relatively inelastic at 0.15 and (-0.13), respectively, though neither were statistically significant. Cross-price elasticity for processed vegetables was 0.05 while cross-price elasticity for processed fruit was 0.02, though neither was statistically significantly different from zero.

Andreyeva, Long, and Brownell (2010) completed an inventory of all empirical US-based food demand work over the past 70 years and found demand of fruit & vegetables have mean own-price elasticities of 0.70 and 0.58, respectively. Consumers are historically most price sensitive to purchases of food away from home (elasticity of 0.81) while they are least price sensitive to egg purchases (elasticity of 0.27). Estimated parameters for vegetables are historically consistent, suggesting that consumer price sensitivity with respect to vegetable purchases has changed very little in the past 70 years. Fruit demand, on the other hand exhibits a greater range of elasticity parameter estimates. Durham and Eales (2010) inventoried 16 studies on food demand in the US and found that while studies based on market-level data rendered relatively price inelastic fruit demand, those using detailed retail-level data exhibited more elastic fruit demand. In a pedagogical article on food demand, Herrmann and Roeder (1998) support the need for disaggregated consumption data in estimating consumer food demand. They argue that because food retailers in industrialized nations compete with one another on a local level by adopting active pricing strategies, the point-of-sale consumption will be more price elastic compared to the aggregate-level market demand.

Durham and Eales (2010) gathered price and sales data for several fruits from two supermarkets in Portland, Oregon and found all fruits were own-price elastic excepting bananas, but not significantly so. Their sample consisted of individuals across a wide income distribution, which supports the generality of their findings. A 1995 study by Hoch et al. (1995) of Dominick's Finer Foods, a major Chicago-based supermarket chain, found food and beverage categories have own-price elasticities above unity. These authors analyzed 160 weeks of scanner data for

---

<sup>5</sup> USDA ERS, Commodity & Food Elasticities database. July 2012. <http://www.ers.usda.gov/data-products/commodity-and-food-elasticities/demand-elasticities-from-literature.aspx>. 28 Feb 13.

18 product categories (including twelve food categories) at 83 store locations. These results suggest price-elastic consumer food purchasing behavior at the retail level.

The household production process (Becker 1965) implies that higher income households will be less price-sensitive due to their relatively high searching costs compared to low-income households. However, to be responsive to prices, consumers have to be aware of the distribution of prices, which requires considerable time and psychological effort. Households with vehicle access and storage space, thus higher income households, may have greater ability to take advantage of price deals as well as to stockpile (Hoch, Kim, Montgomery, and Rossi 1995). Based on this literature as well as the aforementioned findings of Durham and Eales (2010) and Herrmann and Roeder (1998), I can expect that my disaggregated sample data on fruit & vegetable purchases will exhibit relatively high elasticity even for higher-income households.

#### **2.3.4 Review of Fruit & Vegetable Demand and Income**

Numerous studies have analyzed income constraints on fruit & vegetable consumption. While low-income households generally spend a lower share of their budget on fruits & vegetables than higher income-households, the literature is unclear as to how sensitive fruit & vegetable expenditures are to changes in income or prices. Raper, Wanzala, and Nayga (2002) compared demand elasticities for nine food groups between low-income and higher-income households. Their data consisted of 3,577 household food expenditure diaries and demographics from the 1992 Bureau of Labor Statistics Consumer Expenditure Survey (BLS CES). Fruit & vegetable purchases were equally sensitive to income changes, but more elastic to own-price changes for low-income households compared to higher-income households. Blisard, Stewart, and Jolliffe (2004) examined fruit & vegetable demand behavior across 5,000 households using the BLS CES from 1991 and 2000. These authors compared the expenditure distributions on fruits & vegetables between low-income households (below 130 percent of the poverty line) and higher income households using a stochastic dominance test and found that per capita fruit & vegetable expenditures of higher income households are always greater than those of the lower income households. Their finding was robust for all fresh and processed fruits & vegetables. Further, low-income households were more likely to spend an additional dollar of income on food groups other than fruits & vegetables or to other non-food products compared to higher-income households. While quality differentials on the price margin of the fruits & vegetables could explain the expenditure differences between low and higher-income households, the authors note a 1995 study by Krebs-Smith that found lower income households consume smaller quantities of fruits & vegetables.

Studies by Wilde, McNamara, and Ranney (1999 and 2000) found that household recipients of food stamps are more likely to consume meats, added sugars, and fats than they would without the program, and receipt of food stamps are not associated with higher levels of fruit & vegetable consumption. These studies suggest lower

income households consume less fruits & vegetables in aggregate, but are also less likely to substitute fruits & vegetables for non-fruits & vegetable foods if the relative prices of fruits & vegetables fall, or if they are subsidized through food stamps. Thus, for low-income households, fruit & vegetables appear to have very low price elasticity of demand.

### **2.3.5 Review of Fruit & Vegetable Demand and Food Geography**

While income has a significant effect on food consumption behavior, food access policy aims to influence consumption behavior among low-income households through the food environment. Previous research suggests proximity, in addition to income, is correlated with fruit & vegetable consumption. Using retail-level receipt data from a non-profit produce grocer, Weatherspoon et al. (2013) found income elasticities to be significantly more elastic for residents of a Detroit food desert than for average American consumers. In other words, income played a significant role in determining the purchasing behavior of consumers. Their study found purchases of fruits would increase by over one percent for a one percent increase in income, suggesting that increasing household income, or lowering fresh fruit prices can play a substantial role in increasing fresh fruit consumption of food desert residents. Their study further suggests that residents of a Detroit food desert exhibit significantly different fruit purchasing behavior than the average American.

Ver Ploeg et al. (2009) examined food purchases among food stamp recipients from June 1996 through January 1997 to measure relationships between supermarket access and household purchases of perishable and non-perishable fruits, vegetables, and milk. Results of their Tobit censored regressions revealed households that did not shop mainly at a supermarket purchased significantly less perishable vegetables and fruits compared to households that shopped frequently at supermarkets. The study, however, lacked geographic data on food retail access. The researchers relied on respondent's stated shopping frequency and the type of store at which they most often shopped to proxy for food retail access.

Results from the 1996-1997 National Food Stamp Program Survey by Rose and Richards (2003) showed that food stamp recipient households living greater than five miles from their principal food store consumed 63 grams of fruit and 36 grams of vegetables less per day than those living within one mile of their principal food store. These measurements translate into approximately 0.8 servings of fruit and 0.5 servings of vegetables less per day for food stamp recipients living over five miles from their main food store.

Researchers in the UK completed one of the earliest "before and after" studies on food access and health in Leeds, England in 2002. In 2000, the British supercenter Tesco opened a 70,000 square foot retail store in the highly deprived local authority housing estate area, Seacroft on the edge of Leeds. Wrigley, Warm, Margetts, and Whelan (2002) reviewed resident survey data on food consumption before and after the opening of Tesco. Respondents

with the lowest initial levels of fruit & vegetable consumption (less than one portion per day) increased their fruit & vegetable consumption from an average of 4.1 portions per week to 9.8 portions per week. A broader analysis on respondents who initially consumed less than two portions of fruits & vegetables per day showed similar results: average fruit & vegetable consumption for this group increased from 9.1 portions per week to 12.3 portions per week. The UK study lends convincing evidence to the impact access can have on improving diet. US-based research, however, remains inconclusive.

## **2.4 Contributions**

The methods developed in this thesis introduce preferences into the identification of constraining food geographies. I specifically contribute to the current literature on food policy and access through the following: First, I utilize first-hand survey data on a comprehensive set of fresh food purchases and observed store-specific price data to infer fresh produce demand and consumption behavior. Second, while previous food access literature largely focuses on consumption differences between low and high income households, I attempt to identify the effect of food geography on consumption independent of income levels or significant demographic variation. The respondent sample under review is comprised of a homogenous cross-section of upper middle-class, educated households thus I am confident that any behavioral variation correlated to food geography is either a result of the exogenous environment itself, or personal preferences unrelated to income constraints or highly variant demographics. Third, by combining a stated preference analysis with a demand system, I seek to detach individual preferences from observed consumption behavior in order to isolate the extent to which proximity can constrain consumption of fresh produce. Thereby, I introduce a methodology that can address causality issues pervasive in pre-existing literature. Finally, previous research on food access has largely ignored produce quality as an important determinant of consumption behavior. I explicitly account for produce quality variation to test how lack of access to high quality produce can inhibit fresh produce consumption.

## Chapter Three: Methodology & Theoretical Framework

### 3.1 Introduction

This thesis asks whether food geography constrains consumption of fresh produce. To empirically test my hypothesis I explore the relationship between food geography and behavior using two methods: The first method analyzes current food consumption behavior and asks how price sensitivity to fresh produce as well as the propensity to substitute between fresh produce and other foods relates to proximity to grocery stores. The second method tests how consumer willingness-to-pay for fresh produce changes based on accessibility of the produce, the level of its quality, and the type of store selling the produce, and how these changes are informed by proximity to food stores.

Using these methods, I explore whether food deserts have constraining effects under the common geographic USDA classification as well as a stricter preference-based classification. Under the geographic definition, I compare results between “1 mile or less” respondents and “Over 1 mile” respondents in order to test whether food geography has any effect on food consumption. Under the preference-based definition, I again compare results between two subsamples, but I refine these subsamples to include those that are least and most likely to be constrained by their food geography. The first subsample of those least-likely to be constrained is comprised of respondents who live within one mile of a grocery store also explicitly reveal that they prefer their current grocery shopping scenario. I refer to this group as the “Unconstrained” group. The second subsample, by contrast, is comprised of respondents most likely to be constrained because they not only live over one mile of a grocery, but also reveal explicitly that they do not prefer their current grocery shopping scenario. I refer to the latter group as the “Constrained” group. This preference-based classification refines the USDA geographic classification to a smaller sample of respondents who reveal that they are either satisfied or unsatisfied with their current food geography. The dissatisfied respondents are clearly constrained by their food geography because they meet the USDA definition of “food desert” residents with respect to their geographic proximity, and further they explicitly state their desire for a new situation. The preference-based classification restricts any statistical noise that may arise by removing those respondents who appear ambivalent about their current food geography. I expect the preference-based classification to reveal differences in consumption behavior if food geography, indeed, can effect food consumption.

As an additional robustness check on constraining effects of food geography, I explore the behavior of respondents who live within one mile of a grocery store, but choose to grocery shop at a store over one mile from their residence.<sup>6</sup> I refer to this group of respondents as those who “Opt to Travel.” Such respondents represent a

---

<sup>6</sup> These respondents reported a travel distance to their main grocery store greater than five minutes. The five-minute cut-off travel time is conservatively based on a 12 mph average driving time. The data verifies the validity of this assumption because

population of individuals who might display identical behavior and preferences regardless of their proximity to a grocery store. I compare behavior and preferences of these respondents who willingly travel, to the “Constrained” respondents who must travel over one mile to access a grocery store in order to test how necessitated travel as opposed to willing travel can affect consumption behavior. The following sections present the methodology used to implement these analyses.

### **3.2 Demand System: “The Almost Ideal Demand System”**

A system of demand for grocery products provides a method for comparing consumption behavior between consumer groups. The food demand model for this thesis relies on a system-wide estimation, where I approach fruit & vegetable demand as simultaneously dependent upon demand for all other groceries. The theory of multistage-stage budgeting and the related concept of weak separability allow me to estimate grocery demand as a function of household grocery expenditures alone (Deaton and Muellbauer 1980b).

Before determining the structure of the food demand system, one must determine how budget shares are affected by income. Demand system specification is grounded in identifying the relationship between total expenditure and budget shares for each good in the demand system. When plotted on a graph, the relationships between total expenditure and budget shares are known as Engel curves. I analyzed Engel curves using linear, quadratic, and locally weighted non-parametric regressions of food budget shares on the log of total weekly grocery expenditures to distinguish between linear (rank 2) and quadratic (rank 3) functional forms of the demand system.

For purposes of specifying the appropriate functional form, I aggregated food expenditures into fresh fruit, fresh vegetables, frozen & canned fruit, frozen & canned vegetables, and other foods to evaluate the impact of grocery store access on non-perishable produce consumption. Ver Ploeg et al. (2009) similarly disaggregate fruits and vegetables into canned and non-canned to test whether households with limited access to supermarkets tend to spend proportionally less of their food budget on perishable foods. The linear and quadratic Engel Curves for these five food categories are presented in Figure 2. Based on the results of the regression lines, there does not appear to be significant deviation from linear behavior, suggesting demand for all food commodities can be represented by rank 2, or linear demand. I disaggregated the non-perishable fruit & vegetable categories into frozen fruit, canned fruit, and frozen vegetables, and canned vegetables, for a total of seven share equations in the final demand system. Engel curves for each of these disaggregated groups behave identically when separated as when combined. However, I separate the non-perishable groups in order to test for differential substitution behavior between fresh and frozen, and fresh and canned fruits & vegetables.

---

of those respondents who provided the name of their neighborhood and who reported travel time less than five minutes, 89.5 percent indeed live within one mile of the nearest grocery store.



Linear demand estimation follows the AIDS specification of Deaton & Muellbauer (1980a):

$$w_n = \alpha_n + \sum_j \gamma_{nj} \log p_j + \beta_n \log\left(\frac{x}{p}\right) \quad (1)$$

where the price index  $P$  in log form is given by:

$$\log P = \alpha_0 + \sum_k \alpha_k \log p_k + \frac{1}{2} \sum_j \sum_k \gamma_{kj} \log p_k \log p_j \quad (2)$$

Budget shares are denoted by  $w_n$  where  $n$  is the index of food categories 1 through 7. Total food expenditure is denoted by  $x$  and price of food categories is denoted by  $p_j$  where  $j$  is a free index on each food category 1 through 7. The  $\alpha_n$ ,  $\beta_n$  and  $\gamma_{nj}$  terms are parameters that satisfy the following linear demand restrictions:

$$\text{Adding-up} \quad \sum_n \alpha_n = 1 \quad \text{and} \quad \sum_k \gamma_{kj} = \sum_j \beta_j = 0$$

$$\text{Homogeneity} \quad \sum_j \gamma_{kj} = 0$$

$$\text{Symmetry} \quad \gamma_{jk} = \gamma_{kj}$$

The adding-up restrictions force total expenditure to be equal to the sum of individual expenditures for each food category. Homogeneity of degree zero means that if all prices and income both double, demand does not change, ceteris paribus. The symmetry condition refers to the Slutsky symmetry requirement of a utility-maximizing demand system. I have set the  $\alpha_0$  term equal to 3, the minimum log expenditure observed in my sample data (Banks, Blundell, and Lewbell 1997).

The demand system parameters  $\alpha_n$ ,  $\beta_n$  and  $\gamma_{nj}$  were estimated in Stata 11<sup>7</sup> by applying nonlinear seemingly unrelated regression to the system of share equations. The adding-up restrictions result in singularity of the error covariance matrix, thus I dropped the seventh “other foods” share equation for estimation purposes to avoid covariance matrix singularity. The estimator is a feasible nonlinear generalized least squares estimator executed using the Stata command `ns1ur`.

Elasticity measures result from the share equations and parameter estimates as follows:

Income elasticity:

$$\eta_n = (\beta_n + w_n) / w_n \quad (3)$$

Marshallian “uncompensated” cross price elasticity:

$$\varepsilon_{nj} = (\gamma_{nj} - \beta_n \alpha_j - \beta_n (\sum_j \gamma_{jj} \ln p_j)) / w_n \quad (4)$$

Hicksian “compensated” cross-price elasticity:

$$s_{nj} = \varepsilon_{nj} + (\eta_n * w_j) \quad (5)$$

---

<sup>7</sup> StataCorp. 2009. *Stata Statistical Software: Release 11*. College Station, TX: StataCorp LP.

Elasticities were calculated at sample medians. Because Marshallian cross-price elasticity accounts for substitution effects as well as income effects resulting from a good's price change, I use the Marshallian elasticity to analyze substitution behavior resulting from price and cost variation of fresh produce.

Prices  $p$  for each food category were calculated based off of an average representative basket of products as follows: For the fresh fruit & vegetable categories as well as frozen and canned fruit & vegetable categories, I calculated the average amount of fruits & vegetables purchased per week across respondents (e.g. one pound of bananas, two pounds of apples, one pound of carrots, one bag of frozen green beans). Respondents provided their grocery purchase history in terms of the week or month prior to the survey depending on which time frame was easier for them. To calculate average weekly fruit & vegetable purchases, I normalized all responses into weekly quantity terms. I then calculated the price of these average quantities of fresh fruits, fresh vegetables, frozen and canned fruit, and frozen and canned vegetables at each major grocery store in Champaign County using observed contemporaneous price data from these stores. Each respondent provided information on the allocation of their grocery expenditures across all food retailers (e.g. respondent A spent 10 percent of their food expenditures at Strawberry Fields, 50 percent at Schnucks, and 40 percent at Meijer). Based on these allocations, I calculated a price index for each food category for each respondent. Respondents of my survey were asked specific questions on quantity purchased only for fruits & vegetables. Thus, the quantity data I used to generate a price index for the "other foods" category is based on average weekly food consumption among adults 19 to 50 years of age from the USDA's 2001-2002 National Health and Nutrition Examination Survey (NHANES). NHANES data provides average weekly quantities for grains including snacks, cereals, and breads; beverages; dairy products; proteins; and other food items such as condiments, oils & fats, candy, frozen entrees, etc. (Carlson, Lino, Juan, Hanson, and Basiotis 2007). Prices for the "other foods" category are also based on observed contemporaneous price data for each major grocery store in Champaign County. The resulting data set of respondents who provided information on their grocery purchases consisted of 886 observations for all variables in equation (1).

### **3.2.1 Expected Results of Demand Model Analysis**

One mechanism by which lack of access to grocery stores could inhibit people from consuming fresh produce is through limiting the frequency of shopping trips to a distant store. If this mechanism holds, I expect respondents who live far from a grocery store to exhibit both income- and price-sensitivity to fresh produce because their opportunity cost of purchasing perishable foods will be higher relative to respondents who live near to a grocery store and can grocery shop more frequently. Income- and price-sensitivity will manifest through a relatively elastic income- and price- elasticity of demand. Conversely, I expect these respondents to have relatively inelastic income- and price- elasticities for non-perishable foods if non-perishables provide a basis for their diet. This hypothesis follows the findings of Weatherspoon et al. (2013) who found "food desert" residents had a high

income elasticity of demand for fresh fruit; as well as Ver Ploeg et al. (2009) who found households who grocery shop infrequently spend a greater share of their grocery bill on non-perishable foods.

Fresh produce share categories may exhibit higher income- or price- elasticities because the respondent prefers non-perishable goods to fresh produce, and treats the fresh produce consumption as low priority, residual consumption. However, if a respondent would prefer to purchase fresh produce, but is constrained by their distance to a grocery store, I expect that they will substitute canned or frozen fruits & vegetables for fresh produce to a greater extent than respondents who live near to a grocery store. In other words, the cross-price elasticity of demand between fresh vegetables (fresh fruit) and canned or frozen vegetables (canned or frozen fruit) will be positive and relatively elastic, indicating that if the price of fresh produce increases, the demand for non-fresh produce increases. Such behavior would imply that the extra distance to a grocery store forces the respondent to compromise their desired consumption of fresh produce for non-fresh produce. This hypothesis follows the findings of Ver Ploeg et al. (2009), Rose and Richards (2003), and Wrigley et al. (2002) who each found increased distance from a grocery store or infrequent grocery shopping is correlated with lower consumption of fresh produce.

As a robustness check on the effect of travel time on consumption, I compare the behavior of those who willingly travel far, the “Opt to Travel” group, and those who are forced to travel far, the “Constrained” group. Both groups should exhibit similar price-, income-, and cross-price elasticities if travel time to a grocery store can directly alter consumption behavior of fresh produce.

In short, my hypotheses for the demand estimation are as follows:

**Hypothesis 1:** Respondents who live over one mile from a grocery store exhibit more elastic demand for fresh produce and more inelastic demand for “other foods” relative to respondents who live within one mile of a grocery store.

**Hypothesis 2:** Respondents who live over one mile from a grocery store are more likely to substitute fresh for non-fresh produce, especially those who would like to consumer fresh produce, but are constrained by their food geography.

**Hypothesis 3:** If travel time to a grocery store can directly alter consumption behavior of fresh produce, those who “Opt to Travel”:

- a) Are similarly price sensitive to fresh produce as “Constrained” respondents, and
- b) Exhibit similar substitution behavior between fresh and non-fresh produce as “Constrained” respondents.

The demand system will illustrate how food geography, alone, may compromise fresh fruit & vegetable consumption. I anticipate this effect will be similar for all respondents living far from a grocery store relative to

those living near to a grocery store. However, results of the choice experiment will provide some indication as to how each of these groups “feel” about those compromises, which I expect will demonstrate that the “Constrained” group are significantly less satisfied with the compromise than are the “Over 1 mile” group.

### **3.3 Choice Experiment**

#### **3.3.1 Theoretical Framework**

Grocery shopping involves several decisions including where to shop, what to buy, and how much to spend. Because grocery stores are “lumpy goods” these decisions are constrained not only by budget, but also by the food retail options available, particularly in “food deserts” where the array of food stores is limited. To illustrate, a consumer living in downtown Urbana with \$20 to spend can choose to shop for apples at Store A, which is 15 minutes away, Store B, which is eight minutes away, or Store C, which is three minutes away. Figure 3 provides a graphical illustration. Each of these stores offers the consumer a unique combination of the quantity of apples they can buy with \$20 based on the apple price, and time spent traveling to the store, or, viewed differently, time saved. The consumer must decide where to shop based off of their inherent value of time and desire for monetary savings. Simultaneously, the consumer must make a tradeoff between the quantity of apples they can buy and the quality of the apples. For example, assume Store C has the highest quality apples, but the most expensive, whereas Store A has the lowest quality apple, but the least expensive. The consumer can purchase more apples at Store A, but must forgo quality for higher quantity if they choose to shop there. See Figure 4 as an example. The consumer must also consider the travel time required to get to Store C, which may be a short trip versus Store A, which may be a longer trip on the expressway. Further, the consumer must decide on the type of store they shop at, taking into consideration size, ease of navigating the store, the amount of non-food items and services available, and other factors such as ownership structure. Store C has the most expensive produce and is also the smallest store with few services and non-food items, while Store A is the largest option with a wide array of non-food items and services. See Figure 5 as an example.

In each of these decisions, the consumer faces a discrete choice among the food retail opportunity set. “Food desert” residents are theoretically forced to trade off distance for quality, price, store amenities, and several other factors to an extreme. I investigate whether the obligation to make such tradeoffs has any effect on produce consumption. Such tradeoffs are impossible to observe from individual behavior because the consumer’s choice of grocery store is predicated on their internal, unobservable rankings of importance among distance, quality, price, and several other factors. Consequently, I rely on a stated preference approach to parse consumers’ partiality across these factors.

Stated preference methodologies enable economists to estimate welfare changes that are usually hypothetical and cannot be observed in revealed behavior. This thesis utilizes a type of stated preference method known as a choice experiment to evaluate people’s willingness- to-pay for various grocery store attributes. A choice

experiment provides the framework to simulate how survey respondents choose among a set of grocery store alternatives. In a typical choice experiment the decision maker, or respondent, faces a question asking them to choose between several alternatives with differing characteristics. The respondent decides which alternative they most prefer. The respondent answers a series of such questions, whereby the attributes of each alternative differ slightly according to the experimental design. By varying the attribute levels for each question, the researcher can estimate the respondent's willingness to substitute between attributes. The consumer's choice is then modeled in probabilistic terms (Greene (a)). Monetary cost can be used as one of the attributes to calculate the implicit price, or willingness- to-pay, for a change in a particular attribute. Such calculations also enable the researcher to calculate a respondent's willingness- to-pay to move from their current situation to a given choice scenario. The framework of a choice experiment needs to be realistic enough to explain an individual's previous choices, and should be capable of assessing the support for alternatives not presented in the choice set. The choice set generation should have mutually exclusive, exhaustive, and finite alternatives. Consequently, the experiment designer must also account for the range of alternatives and influences affecting the decision to choose one alternative over another, such as demographic variables (Hensher, Rose, and Greene 2005; Greene (b)). The designer must also set levels within each attribute such that the levels may imply meaningful changes in utility for changes in the attribute (Alpizar, Carlsson, and Martinsson 2001).

### **3.3.2 Survey Design**

I developed a survey to measure the value of food access and identify preferential and environmental factors determining fruit & vegetable consumption. The survey asked respondents several questions pertaining to their grocery shopping activities from the month prior to the survey, roughly December 2012 to early January 2013, as well as their typical food consumption and purchasing routines. The choice experiment component presented respondents with six sets of hypothetical grocery stores and asked respondents to indicate their preferred shopping scenario based on variation in these grocery store attributes: store type, travel distance to the store, variety and quality of produce, and price of one week's worth of groceries from that store. Respondents could also opt to choose neither hypothetical grocery store, and instead choose their current grocery store as their preferred shopping scenario. At the start of the choice experiment, respondents were instructed to select their preferred grocery store based off of where they would routinely shop for groceries. Figure 6 provides the varied levels of each of the four grocery store attributes. Figure 7 provides an example of one of the choice questions.

The travel time and produce attributes were motivated by my intent to measure a tradeoff between travel time, produce quality, and variety. I left the definition of "quality" in the produce quality and variety attribute undefined intentionally to ensure that quality perception remained individual-specific across respondents. Outside of the choice experiment questions, respondents provided information on their own definition of produce quality by ranking the relative importance of taste, freshness, growing method, seasonality, and shelf life. I varied the travel time attribute enough to ensure that respondents would perceive an increase from five minutes to 15 minutes as a

significant increase in searching costs, which Hoch et al. (1995) define as a function of the physical distance and driving time between stores. Other attributes I identified through informal interviews with residents of Champaign County and a focus group involving graduate students of the University of Illinois, as well as examining literature on grocery store preferences and determinants of store choice (see Section 2.2.4). I tested the online survey with graduate students of the University of Illinois on multiple occasions and modified the final survey in response to test-takers' feedback. I designed the online survey through the survey platform Qualtrics.<sup>8</sup> The Division of Management Information at the University of Illinois distributed the survey via email to 5,000 faculty and staff of the University of Illinois at Urbana-Champaign. Out of 5,000 individuals who received the survey, 991 provided answers resulting in a 19.8 percent response rate.

To decide the exact ordering and combination of attributes and levels presented to each respondent, I followed standard practice to create a fractional factorial experimental design (Kuhfeld 2010; Louviere et al. 2000; Hensher et al. 2005). Such a design ensures that variations of choices presented to respondents are uncorrelated. I used Kuhfeld's SAS macro (Kuhfeld 2010) to generate a design that achieves a 100 percent D-efficiency and can be implemented with a subset of 36 profiles.<sup>9</sup> I blocked the 36 choice profiles into three sets, resulting in three unique surveys. Each survey asked respondents six choice questions where each choice question consisted of two choice profiles plus a status quo option. The status quo option represented the respondent's current main grocery store. The SAS code used to generate my design as well as the SAS output are presented in Appendix A. Of 991 survey respondents, 863 subjects completed the choice experiment, each of which answered six choice sets with three alternatives. In choice experiments, the unit of observation is each unique choice alternative as opposed to each respondent (Kuhfeld 2010 p. 292). Thus the surveys provided a total of  $863 \times 6 \times 3 = 15,534$  observations for the econometric analysis.

### 3.3.3 Econometric Methods

Econometricians often rely upon structural models of behavior called random utility maximization models to model and interpret discrete choice. Random utility maximization theory asserts that various characteristics of a good endow individuals with utility along with a random component, which captures unobserved differences. The utility  $U_j$  associated with alternative  $j$  is comprised of a systematic, representative utility  $V_j$  and a stochastic, unobservable element of utility  $\varepsilon_j$  as follows:

$$U_j = V_j(x_j, \beta) + \varepsilon_j \quad (6)$$

<sup>8</sup> Qualtrics 2013. Provo, Utah. Version 37,892. <www.qualtrics.com>

<sup>9</sup> A 100% D-efficiency indicates a design is balanced and orthogonal. A design is balanced when each level occurs in the design equally often within each attribute. A design is orthogonal when every pair of levels occurs equally often across all pairs of attributes. Stated more generally, a design is orthogonal when the frequencies for level pairs are proportional or equal (Kuhfeld 2010 p. 58).

where  $x_j$  is a vector of observable attributes of alternative  $j$  and  $\beta$  is a vector of parameters based on observable attributes of the decision-maker. Assumptions on distribution of the unobserved portion of utility determine what type of discrete choice model is used. The multinomial (MNL) and conditional logit (CL) models are widely used for discrete dependent variable analysis because their resulting choice probabilities take on a convenient form (Heiss 2002).

Random utility maximization theory assumes that consumers choose the alternative that endows them with the highest utility (Greene (b)). Suppose we may assign a utility level  $U_{ij}$  to each alternative  $j = 1, \dots, J$  for each decision-maker  $i = 1, \dots, I$ . The utility levels are determined by characteristics of the choice alternatives as well as characteristics of the decision-maker, some of which are identifiable and some of which are unknown. The MNL model is a special case of a CL model in which all explanatory variables are assumed to be individual-specific. In a CL model, a choice among alternatives is treated as a function of the characteristics of the alternatives as opposed to a function of characteristics of the individual decision-makers, thus each  $\beta$  term will be identical across respondents (Kuhfeld 2010). If we consider each choice situation for each respondent, the utility gained by person  $i$  for alternative  $j$  is given by

$$U_{ij} = \beta X_{ij} + \varepsilon_{ij} \quad (7)$$

where  $X_{ij}$  is a vector of attributes of alternative  $j$  relative to individual  $i$  and  $\beta$ , a vector of parameters, acts as a scalar relating the relationship between individual  $i$  and the alternative  $X_{ij}$  to the individual's utility for the alternative. The parameters of vector  $\beta$  are not alternative specific, thus are not subscripted by  $j$ .

The MNL model, however, allows the effect of each independent variable (i.e. attributes of the decision-maker) to vary across choices. The basic MNL utility equation of person  $i$  for alternative  $j$  is given by:

$$U_{ij} = z_i \gamma_j + \varepsilon_{ij} \quad (8)$$

The vector  $\gamma_j$  of alternative-specific parameters relate the characteristics of a respondent to the respondent's utility for the  $j$ th choice. The parameter  $z_i$  is a matrix of individual characteristics that are independent of the alternatives available. While  $X_{ij}$  from the CL model is a vector of alternative-specific characteristics,  $z_i \gamma_j$  from the MNL model is a vector of individual-specific characteristics relating the characteristics of a respondent to the respondent's utility for the  $j$ th choice.

The mixed multinomial logit model (MMNL) combines the MNL and CL models so that the effect of individual characteristics as well as choice characteristics on the probability of choosing an alternative may both be identified simultaneously. The MMNL model assumes that the unobserved portions of utility may follow any specified distribution thus relaxing the CL assumption that respondents are homogeneous with regards to their preferences (Dissanayake 2011). The probability  $P_{ij}$  that an individual  $i$  will choose alternative  $j$  from the choice set of  $m$  total

alternatives does not have a closed form solution, and must be evaluated numerically through iterative simulations. Utility gained by person  $i$  for alternative  $j$  takes the form

$$U_{ij} = z_i\gamma_j + \beta X_{ij} + \varepsilon_{ij} \quad (9)$$

where  $z_i$  is a vector of individual-specific independent variables relevant for individual  $i$  and  $\gamma_j$  is a vector of alternative-specific parameters. Together, these parameters relate the characteristics of a respondent to the respondent's utility for the  $j$ th alternative such that the effect of the independent variables will vary across all choices (Green (b)). The MNL model assumes that the unobserved portions of utility  $\varepsilon_{ij}$  are independently and identically distributed. Thus, the unobserved portion of utility for one alternative is independent of the unobserved portion of utility for other alternatives. McFadden (1974) demonstrated that under these assumptions, the probability  $P_{ij}$  that an individual  $i$  will choose alternative  $j$  from the choice set of  $m$  total alternatives has a closed form solution given by

$$P_{ij} = \frac{\exp(U_{ij})}{\sum_{k=1}^m \exp(U_{ik})} = \frac{\exp(z_i\gamma_j + x_{ij}\beta)}{\sum_{k=1}^m \exp(z_i\gamma_k + x_{ik}\beta)} \quad (10)$$

The unknown parameter  $\beta$  can be estimated by the maximum likelihood procedure whereby one finds the vector  $\hat{\beta}$  that maximizes the joint probability of observing the observed choice outcomes of all consumers in the sample (Heiss 2002; Arnold et al. 1983).

The coefficient estimates for the CL/MNL or MMNL models alone are difficult to interpret directly. Instead, the ratio between the non-monetary parameters and the cost parameter provide the marginal willingness- to-pay (WTP) for a change in each attribute  $\eta$ . The average marginal WTP is given by

$$MWTP_{\eta} = - \frac{\beta_{\eta}}{\beta_{bill}} \quad (11)$$

### 3.3.4 Econometric Specification

I use an MMNL model to estimate parameters from the choice experiment data. Parameter estimates of the MMNL model are sufficiently similar to the CL model, however I employ the MMNL in order to account for preference heterogeneity across respondents and for possible correlations between each individual's survey responses. I compare results of the CL model to that of the MMNL model to ensure that the added flexibility of the MMNL does not significantly alter the parameter estimates. My model specification employs main effect terms as well as interaction effect terms. The "main effects" refer to the four grocery store attributes in my choice experiment, including cost of a week's worth of groceries, travel time required to get to the store, store type, and the produce selection. The interaction terms account for potential non-linear relationships between respondent income, proximity, or travel time and their willingness-to-pay for grocery store attributes.



The econometric specification is given by:

$$U_{ij} = \alpha_j + \beta_1 X_{iBill} + \beta_2 X_{iTravel} + \beta_3 X_{iSpecialty} + \beta_4 X_{iGrocery} + \beta_5 X_{iHQProduce} + \beta_6 X_{iAQProduce} + \beta_7 X_{iBill} * \gamma_i CurrentTravel + \beta_8 X_{iBill} * \gamma_i Distance + \beta_9 X_{iHQProduce} * \gamma_i Income + \beta_{10} X_{iAQProduce} * \gamma_i Income + \varepsilon_{ij} \quad (12)$$

The  $\alpha_j$  term is an alternative specific constant (ASC). This ASC term captures the average effect on utility of all factors that are not included in the model (Greene (b)). The ASC term is a dummy variable equal to 1 if the respondent selected their current situation status quo option, and equal to 0 if the respondent chose a hypothetical option A or B. To account for the potential tendency for people to choose their current store I include an ASC to capture attributes affiliated with the status quo choice that I did not explicitly account for in the experimental design. Tang et al. (2001) found in a sample of 500 households that knowledge of store layout was a significant source of utility of grocery shopping. Each  $X$  term corresponds with one of four grocery store attributes included in the choice experiment: weekly grocery bill (*Bill*), travel time to the store (*Travel*), store type (*Grocery* and *Specialty*), and produce quality and availability (*HQProduce* and *AQProduce*). The coefficient on each of these main effects provides an estimate of the change in utility given a change in the associated attribute. *Bill* and *Travel* are numeric attributes, thus their coefficients can be interpreted on a continuum as the change in utility given a \$1 change or 1-minute change in the attribute, respectively. On the other hand, store type and produce selection are ordinal, discrete attributes. Consequently, I disaggregated the store type attribute into *Grocery* and *Specialty* variables. The *Grocery* and *Specialty* terms are dummy variables that will both equal 0 if the respondent chose the base case of a supercenter. Coefficients on *Grocery* and *Specialty* provide the change in utility from shopping at a grocery store or specialty store instead of the base case of a supercenter. Similarly, I disaggregated the produce selection attribute into *HQProduce* ("high quality produce") and *AQProduce* ("average quality produce") variables. *HQProduce* and *AQProduce* are dummy variables that will both equal 0 if the respondent chose a store with the base case of low quality produce. The coefficient on these variables indicates the probability of choosing to shop at a store with high quality or average quality produce instead of the base case low quality produce.

Recall, utility cannot be observed directly, but we infer the choice that garners the highest utility based on which choice the respondent selects. The model based on equation (12) includes four  $\gamma$  terms that interact the grocery bill and produce quality attributes with person-specific variables including income level (*Income*), proximity to nearest grocery store (*Distance*), and reported travel time to their principal grocery store (*CurrentTravel*). The error term  $\varepsilon_{ij}$  is independent and identically distributed extreme value type I with a zero mean by construction due to the inclusion of the ASC variable (Greene (b)).

Results of the MMNL yield estimates of the median and standard deviations of each of the coefficients' distributions. If any of the attribute coefficients are significant and positive, this implies that the respondent has a higher probability of choosing an alternative with increased levels for those attributes. The higher probability implies higher marginal utility for increases in those attribute levels. Negative attribute coefficients imply disutility. I expect the *Bill* and *Travel* attributes to be negative, for example, because respondents presumably get disutility from paying extra for groceries or traveling further, *ceteris paribus*. The interaction term  $\beta_7 X_{iBill} * \gamma_i CurrentTravel$  identifies the impact of a respondent's reported travel time to their principal grocery store with their disutility from a higher grocery bill. The interaction term  $\beta_8 X_{iBill} * \gamma_i Distance$  identifies the impact of the distance between the respondent's residence and the nearest grocery store with their disutility from a higher grocery bill. These interaction terms control for potential nonlinearities between respondents' marginal WTP for groceries and the time they take and/or distance they travel to get to the grocery store. For example, respondents who live far from, or must travel several minutes to get to, a grocery store may have a lower marginal WTP for groceries because their extra travel effectively increased the cost of their groceries. On the other hand, these respondents might consider the travel time/distance to be a sunk cost, and therefore be willing to pay more for groceries, *ceteris paribus*. The  $\beta_9 X_{iHQProduce} * \gamma_i Income$  and  $\beta_{10} X_{iAQProduce} * \gamma_i Income$  terms identify the impact of income level on willingness-to-pay for produce quality. Positive and significant coefficients imply respondents with higher income gain higher utility from high quality or average quality produce over low quality produce.

By differentiating the utility equation  $U_{ij}$ , the  $\beta$  terms provide the marginal utilities of each attribute. The ratio between the non-monetary parameters and the cost parameter  $\beta_1$  provide the marginal value of each attribute. For example, the marginal value of shopping at a specialty store instead of a supercenter based on equation (12) is as follows:

$$MWTP_{Specialty} = - \frac{\partial U / \partial X_{Specialty}}{\partial U / \partial X_{Bill}} = - \beta_3 / (\beta_1 + \beta_7 * \gamma_i Distance + \beta_8 * \gamma_i CurrentTravel) \quad (13)$$

The econometric models were estimated using Stata's `clogit` and `mixlogit` commands. Results of the MMNL model are presented and discussed in Chapter Five. Results of the CL model can be found in Appendix B.

### 3.3.5 Expected Results of Choice Experiment Analysis

Both endogenous locational sorting near grocery stores as well as environmental treatment effects can influence respondent preferences for grocery store attributes. If endogenous sorting is the driving influence of consumption behavior differentials across respondent groups, I expect their preferences, and thus their marginal WTP for grocery store attributes to be different across groups. For example, I expect that respondents who live near to a grocery store need compensation (negative marginal WTP) for added travel time if their proximity to a grocery

store is a result of their desire to shorten commutes to the grocery store. The “Over 1 mile” respondents, by contrast, I would expect to have a higher (or less negative) marginal WTP for added travel. Previous literature on sorting supports the possibility that people can choose housing locations based on their preferences for amenities at those locations (Bayer and Timmins 2005; Tiebout 1956). Because greater travel distance to a grocery store increases the opportunity cost of purchasing perishable produce instead of non-perishable food items, I expect the “Over 1 mile” respondents to have a lower marginal WTP for produce quality compared to the “1 mile or less” group. I am agnostic on my prediction of respondent preferences for store type. However, I do expect the “Constrained” group to have a higher marginal WTP for a store type different from that which they typically shop, and the converse to be true for the “Unconstrained” group. For example, if I find that the “Constrained” group mainly shops at supercenters, I would expect their marginal WTP to shop at a grocery or specialty store to be high relative to other groups.

On the other hand, if environmental treatment effects, i.e. the exogenous food geography, are driving consumption behavior across respondents, then I do not expect to see significant differences in their preferences. Further, I would expect their consumption behavior to be inconsistent with their preferences. For example, if “Over 1 mile” respondents exhibit high substitution between fresh and non-fresh produce, and very high marginal WTP for produce quality, this would imply these respondents prefer quality produce, but must compromise fresh for non-fresh produce due to an exogenous constraint. Such a result would provide evidence that geography, independent of preferences, can alter produce consumption. I expect consumption behavior relative to preferences to exhibit greater inconsistencies for the “Constrained” group than the “Over 1 mile” group. Because the “Constrained” respondents are, by definition, dissatisfied with their food environment, endogenous sorting around grocery store locations likely has little effect on their consumption behavior whereas endogenous sorting may be driving some of the results for the “Over 1 mile” group.

Comparing the “Opt to Travel” group to the “Constrained” group will inform whether travel time, produce quality, or some other factor serves to “constrain” the “Constrained” group. For example, I can infer that low access to quality produce constrains the “Constrained” group if their marginal WTP for travel time is similarly low to that of the “Opt to Travel” group, but their marginal WTP for produce quality is significantly higher than that of the “Opt to Travel” group.

Alternatively, if the “Opt to Travel” group has similar preferences and behavior to that of the “Constrained” group, this will suggest that the exogenous “food desert” definition overestimates the constraining effects of food geography because the one-mile cut off fails to account for individuals unaffected by their food environment.

To summarize, my hypotheses for the choice experiment are as follows:

**Hypothesis 4:** If endogenous locational sorting near grocery stores explains divergent consumption behavior between respondents who live less than one mile from a grocery store and respondents who live over one mile from a grocery store:

- a) Preferences between the two groups are dissimilar, and
- b) Preferences are consistent with each group's consumption behavior.

**Hypothesis 5:** If exogenous treatment effects explain divergent consumption behavior between respondents who live less than one mile from a grocery store and respondents who live over one mile from a grocery store:

- a) Preferences between the two groups are similar, and
- b) Consumption behaviors between the two groups are dissimilar.

**Hypothesis 6:** If exogenous treatment effects explain consumption behavior to a greater degree than locational sorting near grocery stores, preferences for the "Constrained" and "Unconstrained" groups will be similar, but consumption behavior of the "Constrained" group and the "Unconstrained" groups will be more contrasting compared to that of the "1 mile or less" group relative to the "Over 1 mile" group.

While the methods presented in this thesis do not provide a comprehensive, statistical means for disentangling the endogenous from exogenous forces impacting food consumption, they do help to inform which of these two forces may have greater influence on food consumption.

### 3.4 Robustness Tests

Because my methodology relies heavily on classifications of respondent groups, I alter the classifications in order to "stress test" the sensitivity of my results to these classification. I alter the classifications in three ways. First, I change the geographic constraint classification from the USDA's 1-mile cut-off to a 2-mile cut-off to test whether results are robust to alternative distance cut-offs. Second, I use only a lower-income sub-sample of respondents to test if the above-average income levels of my respondents bias my conclusions. Third, I expand the sample of "Constrained" respondents to test for any small-sample biases in the preference-based classification analyses. In each of these three robustness checks I mimic the aforementioned methodology where I use a demand system and a choice experiment to compare behavior and preferences across respondent groups.

## Chapter Four: The Data

### 4.1 Data Collection

The following section summarizes key statistical findings from the survey discussed in the previous chapter. In addition to the choice experiment, respondents provided copious information on both their grocery shopping between December 2012 and January 2013, as well as their typical grocery shopping routines. Survey data includes information on where respondents purchased groceries, how often they shopped, the travel time to these stores, their mode of transportation, their average grocery expenditures per household, as well as qualitative questions on produce quality preferences and grocery store preferences. Respondents were asked to provide detailed information on which food types they purchased at each food retailer, and the quantity of fruits & vegetables they purchased. The responses include demographic information, as well as identification of the respondent's local neighborhood in Champaign & Piatt counties. I collected price data contemporaneously from 14 grocery and food retailers in Champaign County so as to estimate individual budget shares spent on staples as well as fruits & vegetables. I also geocoded all food retailer locations in Champaign & Piatt counties and calculated the distance from each respondent's neighborhood to their surrounding food stores. Food retailer locations and industry classifications were sourced from LexisNexis database.

### 4.2 Summary Statistics: Demographics, Food Environment, and Budget Shares

Table 1 compares demographic characteristics of the respondent sample to the population of Champaign County. The sample is biased toward a higher income and higher education level than is representative of Champaign County. Because the marginal effects of grocery store proximity and produce consumption are of greatest interest for this thesis, this bias does not diminish the relevance of my results. Rather, the relative homogeneity of the sample ensures income or education disparities will not spuriously drive any consumption variation across the sample.

**Table 1:** Demographic comparison of Champaign County population and survey sample

Demographic	Sample Average	Champaign County* Average
Income (thousand)	\$87.5	\$59.2
Age (adults over 18)	53	39
Education (% of adults with Bachelor's degree)	99.6%	42.1%
Household size	3	2
Car Access (% of households with)	99.0%	98.8%

\* Source: American Fact Finder 2010 US Census

While car access is ubiquitous across the sample, a minority of respondents reported they use alternative means to travel to the grocery store. Figure 8 illustrates the types of transportation used by the sample. Eighty-five percent of respondents drive to their grocery store, on average. Ten percent ride a public bus, and 4.1 percent walk or bike.

Out of 991 total respondents to my survey, 746 provided some information on their geographic location. Of the 746 respondents, 685 indicated they live in either Champaign or Piatt counties. Finally, of these 685 individuals, 571 provided detailed information on their neighborhood location. Subsequent analyses based on neighborhood location and proximity to food retailers will rely on this subset of 571 individuals.

Figure 9 provides a visual map of food retail and restaurant locations in Champaign & Piatt counties. Figure 10 presents a map of respondent neighborhood locations in Champaign & Piatt counties relative to the location of grocery stores. Respondents to the survey live among 36 distinct neighborhoods in Champaign & Piatt counties.<sup>10</sup> As indicated by the dot density continuum in Figure 10 as well as Table 2, the majority of respondents live in the West Urbana neighborhood of Urbana or the Downtown neighborhood of Champaign.

Distance from respondent's neighborhoods to surrounding food retailers were estimated by generating geographic centroids for each neighborhood and calculating the distance in miles from the centroid to surrounding food retailers. All geocoding, centroid calculation, and near distance calculations were performed using ArcMap 10.0.<sup>11</sup> Nearest distance to grocery stores, supercenters, and convenience stores by neighborhood can be found in Table 2. LexisNexis designated retailers as grocery stores and convenience stores following the Standard Industrial Classification (SIC) system. Because the SIC system does not differentiate supercenters from grocery stores, I designated Wal-Mart, Meijer, Target, and Sam's Club as supercenters. The "convenience store" designation includes drug stores, pharmacies, dollar stores, and gas stations.

The Downtown neighborhood of the city of Champaign claims the largest portion of survey respondents, 13.9 percent of the sample, followed by West Urbana, East Urbana and Clark Park of Champaign. Residents of Savoy face the shortest distance to their nearest grocery store, supercenter, and convenience store. The geographic centroid of Savoy is 0.1 miles from the nearest grocery store and supercenter, and 0.2 miles from the nearest convenience store. Residents of Longview, by contrast, must travel the furthest to get to a grocery store. The geographic centroid of Longview is 15.3 miles from the nearest grocery or supercenter store.

---

<sup>10</sup> Geographic definitions for neighborhoods of Champaign & Piatt counties were sourced from Homeplace Advisors, LLC (<http://yourchampaignhome.com/about-homeplace-advisors/>)

<sup>11</sup> ESRI (Environmental Systems Research Institute). 2010. ArcMap 10.0. ESRI, Redlands, California.

**Table 2: Neighborhood sample population and average distance to nearest food retailer**

No.	Neighborhood	City	Sample Pop.		Census Pop.	Miles to Nearest Store		
			Number	Share		Grocery	Super center	Convenience
1	Downtown	Champaign	114	13.9%	857	0.6	1.7	0.1
2	West Urbana	Urbana	104	12.7%	4,354	0.6	1.8	0.3
3	Clark Park	Champaign	80	9.8%	675	0.9	2.2	0.3
4	East Urbana	Urbana	80	9.8%	1,705	0.4	1.7	0.1
5	Savoy		63	7.7%	7,312	0.1	0.1	0.2
6	Cherry Hills	Champaign	60	7.3%	2,451	1.6	2.3	0.5
7	Mahomet		51	6.2%	10,042	0.3	8.1	0.2
8	Ironwood	Champaign	38	4.6%	3,489	1.8	4.6	1.6
9	Sawgrass	Champaign	26	3.2%	1,118	0.2	3.1	0.8
10	Ponds of Windsor	Champaign	22	2.7%	1,048	0.7	1.6	0.8
11	Boulder Ridge	Champaign	20	2.4%	1,060	0.7	2.6	1.6
12	Campus Town	Champaign	19	2.3%	17,026	0.5	2.6	0.2
13	Monticello		19	2.3%	4,384	0.2	17.3	0.1
14	St. Joseph		17	2.1%	4,407	6.9	6.9	0.3
15	Myra Ridge & South Ridge	Urbana	15	1.8%	872	0.5	0.5	0.9
16	Stone Creek	Urbana	12	1.5%	497	1.0	1.1	1.1
17	Trails at Abbey Fields	Champaign	10	1.2%	38	1.4	4.0	0.7
18	Ashland Park	Champaign	8	1.0%	1,011	0.7	0.7	0.9
19	Savannah Green	Urbana	7	0.9%	3,492	0.5	1.1	0.4
20	Beringer Commons	Urbana	6	0.7%	1,282	0.6	0.6	0.5
21	Tolono		6	0.7%	3,460	4.7	4.7	0.3
22	Rantoul		6	0.7%	13,263	1.2	1.2	0.5
23	Sidney		5	0.6%	1,248	7.0	7.0	3.7
24	Philo		4	0.5%	1,474	5.8	5.8	0.2
25	Pesotum		3	0.4%	557	9.8	9.8	5.2
26	Ogden		3	0.4%	802	10.8	10.8	3.5
27	Longview		3	0.4%	153	15.3	15.3	0.1
28	Ivesdale		3	0.4%	267	8.3	13.2	6.5
29	Fisher		3	0.4%	1,881	8.3	8.9	0.4
30	Thomasboro		2	0.2%	1,127	4.1	4.1	0.4
31	Seymour		2	0.2%	360	5.9	9.2	6.1
32	Rolling Acres	Champaign	2	0.2%	282	2.1	2.2	0.4
33	Hensley		2	0.2%	801	3.3	3.3	3.7
34	Bondville		1	0.1%	414	3.2	6.2	3.1
35	Somer		1	0.1%	1,178	4.6	4.9	4.3
36	Sadorus		1	0.1%	412	8.0	8.0	4.7

Table 3 states the sample's average geographic distances to food retailers by type for the full sample as well as the five subsamples under review. Of all food retailers, convenience stores are highest in number totaling 140 in Champaign & Piatt counties, while supercenters are lowest in number, totaling eight in Champaign & Piatt counties. As expected, the full sample's average distance to supercenters is an extra 2.2 miles further than distance to a grocery store. Distance to a convenience store is half a mile, on average. Respondents in the "Unconstrained" category live closest to all store types while respondents in the "Constrained" category live furthest. The "Status Quo" column in Table 3 refers to the choice experiment survey in which respondents answered six questions asking them to choose between two hypothetical new grocery stores, or their current grocery store, e.g., their "status quo" situation. See Figure 7 as an example of the choice experiment questions. The more often a respondent chose their status quo, the more likely they were happy with their current situation, and preferred

their status quo grocery shopping scenario to any other possible scenario. “Unconstrained” respondents chose the status quo option five times out of six, on average, whereas the “Constrained” respondents chose the status quo option approximately one time out of six, on average.

**Table 3:** Distance (miles) to nearest food retailers by type and 'Status Quo' selection frequency

	Retailer (Population)									Status Quo*	Observations
	Grocery (40)			Supercenter (8)			Convenience (140)				
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max		
Full Sample	1.2	0.1	15	3.4	0.1	17.3	0.5	0.1	6.5	3.3	886
1 mile or less	0.5	0.1	1	2.9	0.1	17.3	0.3	0.1	1.6	3.4	455
Over 1 mile	4.3	1.2	15	5.2	1.2	15.3	1.3	0.1	6.5	3.2	116
Unconstrained	0.5	0.1	0.9	2.5	0.1	17.3	0.2	0.1	1.6	5.0	221
Constrained	4.1	1.0	11	5	1.1	13.2	1.4	0.2	6.5	1.1	56
Opt to Travel	0.5	0.1	1	3.5	0.1	17.3	0.3	0.1	1.6	3.0	260

\*Number of times out of 6 respondent chose their current grocery shopping situation over a new situation

The significant “Opt to Travel” respondent group size suggests close proximity to a grocery store does not guarantee patronage of that grocery store. Of the 571 respondents who provided location information, 80 percent (343 respondents) live less than one mile from a grocery store. Yet, of the 80 percent who live less than one mile from a grocery store, nearly 76 percent (260 respondents) travel over five minutes to get to their main grocery store. See Figure 12 and Figure 13 for a graphical illustration. This means that at least three quarters of the “1 mile or less” sample (with location information) traveled past a grocery store near their house and instead chose to shop at a store further away. In total, the “Opt to Travel” group comprises over 46 percent of the respondents with location information. The 46 percent accounts even for those who traveled to a grocery store to or from work, because the average distance to the nearest grocery store from the center of University of Illinois campus is 0.6 miles - less than one mile.

Table 4 compares grocery shopping habits across respondent groups. Recall, all respondents shopped at a grocery store at least once in the month prior to the survey. Table 4 highlights how many respondents shopped at non-grocery store retailers in addition to a grocery store. “Specialty” food retailers include grocery stores specializing in ethnic and regional cuisine, health or natural foods, specialty produce or meat and cheese products, as well as bakeries. Respondents in the “Unconstrained” group had the highest probability of shopping at a specialty store. Over 65 percent of respondents in the “Unconstrained” group indicated that they had shopped at a specialty store in the month prior to the survey. These respondents also had the least incidence of both supercenter patronage of all four groups, 48.4 percent, and convenience store patronage, 37.1 percent. In contrast, respondents in the “Constrained” group were least likely to shop at a specialty store and most likely to shop at both a supercenter and convenience store compared to all other groups. Over 89 percent of the “Constrained” respondents shopped at a supercenter in the month prior to the survey, and 48.2 percent shopped at a convenience store prior to the survey.



Surprisingly, the “Constrained” group does not appear to be income constrained, as they had the highest mean and median income of all other groups. A comparison of “1 mile or less” and “Over 1 mile” respondents mimics that of the “Unconstrained” and “Constrained” group, but to a lesser degree. Proximity to a grocery store appears to be correlated with a higher likelihood of shopping at specialty stores, and a lower likelihood of shopping at supercenters or convenience stores. Proximity is also correlated with more frequent grocery shopping. The “Unconstrained” group grocery shopped most often in the month prior to the survey, nearly six times whereas the “Opt to Travel” group grocery shopped the least often.

The “Opt to Travel” respondents appear to be cost-conscious, value-driven shoppers. Of the people who opt to travel, 71.2 percent indicated they had shopped at a supercenter in the month prior to the survey, whereas 48.4 percent of the “Unconstrained” group had shopped a supercenter in the past month. This difference was not due to locational convenience: the respondents who opt to travel lived approximately one mile further from a supercenter, on average, than individuals in the “Unconstrained” group (see Table 3).

**Table 4:** Patronage of alternative food retailers, shopping frequency, and income bracket by respondent group

	Incidence of shopping at alternative retailers in past month			Grocery trips per Month*		Income Bracket^		Obs.
	Specialty	Supercenter	Convenience	Mean	Median	Mean	Median	
Full Sample	56.3%	60.3%	40.7%	5.2	4	5.4	5	886
1 mile or less	60.7%	57.6%	40.4%	5.4	5	5.4	5	455
Over 1 mile	52.7%	64.3%	41.8%	5.1	4	5.5	5	423
Unconstrained	65.6%	48.4%	37.1%	5.8	5	5.6	5	221
Constrained	50.0%	89.3%	48.2%	5.3	4	5.9	6	56
Opt to Travel	53.8%	71.2%	46.9%	4.4	4	5.5	5	260

\*Number of times, on average, that the respondent shops at their main grocery store, ie the grocery store where they spent the largest portion of their grocery bill in the month prior to the survey

^Income bracket of 4 denotes annual income between \$50-75K; 5 denotes \$75-100K; 6 denotes \$100-125K

Results of Table 4 suggest that the respondents who are most satisfied with their food geography frequently shop at specialty stores and infrequently shop at supercenter or convenience stores, whereas the opposite is true of respondents who are least satisfied with their food geography.

All respondents of the survey patronized at least one grocery store in the month prior to the survey. No respondent was forced to, nor willingly shopped only at convenience stores. Approximately 40.7 percent of the sample purchased food at a convenience store at least once in the month prior to the survey. Comparing the distance to the nearest grocery between this subset of convenience store-shoppers and the entire sample, convenience store-shoppers on average live further from a grocery store, though the difference is minimal (see

Figure 11). Convenience store-shoppers do not live any closer, on average, to convenience stores than does the entire sample. In fact, convenience store-shoppers live 0.6 miles on average from the nearest convenience store, compared to 0.5 miles for the entire sample. Extra distance to a grocery store may induce individuals to patronize a closer convenience store, but proximity to a convenience store itself does not seem to induce people to shop there.

Table 5 shows the prices, budget shares, zero-response frequency, and total grocery expenditure the sample incurred, on average, for a week's worth of groceries. Appendix C provides a complete list of each fruit & vegetable type for which survey respondents provided purchase information. Respondents could also provide a free response for any fresh, canned, or frozen produce item that I did not ask about in the survey. Grocery purchases were categorized into seven broad categories for purposes of estimating the demand system: fresh vegetables, fresh fruit, canned vegetables, canned fruit, frozen vegetables, frozen fruit, and other foods. The price indices are based on the sample's average quantity purchased per category in one week, and the observed price of these items averaged across 14 grocery retailers. The non-fruit and vegetable "other foods" category is both the most expensive category and makes up the largest share of respondents' grocery bill, followed by fresh vegetables, fresh fruit, canned vegetables, frozen vegetables, and frozen and canned fruit. Respondents purchased "other foods" more often than any other category, since the number of zero entries ("Zero's") is the lowest. Frozen fruit products were purchased least often of all food categories. Average total expenditure on groceries per respondent per week was \$117.6.

**Table 5: Food share and expenditure summary statistics**

	Price Index	Zero's	Budget Share			
			Mean	SD	Min	Max
(1) Fresh Vegetables	\$12.9	97	15.6%	14.4%	0.0%	91.9%
(2) Fresh Fruit	\$10.3	125	11.9%	12.7%	0.0%	75.8%
(3) Canned Vegetables	\$1.4	554	1.6%	3.3%	0.0%	32.1%
(4) Canned Fruit	\$0.4	704	0.5%	1.5%	0.0%	15.5%
(5) Frozen Vegetables	\$1.2	467	1.6%	3.6%	0.0%	46.7%
(6) Frozen Fruit	\$0.4	749	0.5%	1.8%	0.0%	21.5%
(7) Other Foods	\$65.1	23	68.1%	24.5%	0.0%	100.0%
Weekly grocery expenditure			\$117.6	\$67.0	\$10.0	\$500.0
<i>Obs</i>	<i>878</i>					

Table 6 provides correlations between respondents' stated travel time to their main grocery store, the distance to their nearest grocery store, and their grocery expenditure share spent on fresh fruits & vegetables, and frozen and canned non-fresh fruits & vegetables. Travel time and distance have a positive relationship, indicating people that live further from a grocery store travel further. Consumption of fresh fruits & vegetables has a negative relationship with both travel time and distance, while consumption of non-fresh fruits & vegetables has a small, but positive relationship. These correlations provide some evidence that individuals living nearby a grocery store

are more likely to consume fresh fruit & vegetables, instead of non-fresh, however I cannot infer any causal relationship based on this information alone.

**Table 6: Correlation Matrix**

$\rho_{x,y}$	Travel time to store	Distance to store	Budget share of fresh F&V	Budget Share of all non-fresh F&V
Travel time to store	1.000			
Distance to store	0.293	1.000		
Budget share of fresh F&V	-0.046	-0.065	1.000	
Budget share of all non-fresh F&V	0.013	0.055	0.130	1.000

### 4.3 Comparison of Preferences and Budget Shares

My sample provides evidence that proximity to a grocery store has bearing on food purchasing behavior and preferences. Table 7 presents budget shares of each respondent group: “Over 1 mile”, “1 mile or Less”, “Opt to Travel”, “Constrained”, and “Unconstrained.” “Unconstrained” respondents spend the greatest amount of their budget on fresh produce and least on non-fresh fruits & vegetables. The “Unconstrained” respondents spend a full 10.0 percent more of their grocery bill on fresh produce compared to “Constrained.” Similarly, the “1 mile or less” respondents spend a greater portion of their grocery bill on fresh produce and less on non-fresh produce and non-fruit & vegetable foods compared to the “Over 1 mile” group. This suggests that proximity to a grocery store is positively related to consumption of fresh produce relative to non-fresh produce and non-fruit & vegetable foods.

**Table 7: Average grocery expenditure share by respondent group**

	Respondent Group				
	Unconstrained	1 mile or less	Opt to Travel	Over 1 mile	Constrained
Fresh F&V	30.4%	30.1%	29.4%	26.6%	24.4%
Canned F&V	1.8%	2.0%	2.1%	2.3%	2.6%
Frozen F&V	1.8%	2.1%	1.9%	2.2%	2.1%
Non F&V	66.0%	65.8%	66.6%	68.8%	70.8%
Obs.	221	459	264	427	56

Despite these variations in fresh produce consumption, all respondents to my survey spend a significantly greater portion of their food expenditures on fresh produce compared to the national average. Table 8 shows that even among a sub-sample of lower-income respondents, fresh produce consumption is 18.0 percent higher than that of the national average and almost 19.0 percent higher than the national low-income average. Blisard, Stewart, and Joliffe (2004) found that fresh produce consumption is higher among high-income households. Further, graduate education is positively related to high produce consumption independent of income-levels. Thus, fresh produce

consumption among survey respondents is biased upwards due simultaneously to the high education and income levels of my sample.

**Table 8:** Average grocery expenditure share of national and low-income populations

	<b>Lower Income Respondent Group*</b>	<b>National Population^</b>	<b>Low-Income Population^</b>
Fresh F&V	28.6%	10.6%	9.7%
Processed F&V	4.8%	6.8%	6.8%
Non F&V	66.6%	82.6%	83.5%
Obs.	283	5,000	
*Respondents with annual household income less than or equal to \$75,000 per year			
^Economic Research Service, based on the Consumer Expenditure Survey			
Blisard, Stewart, Joliffe (2004)			

The five respondent groups share similar preferences for produce with notable exceptions. Table 9 displays how respondents ranked five characteristics of produce – Freshness, Seasonality, Appearance, Growing Method used to grow the produce, and Shelf Life – based on which have the greatest bearing on their produce purchases. Freshness and Appearance are the first and second most important determinants of produce purchases across all groups. However, respondents with a grocery store over one mile from their neighborhood hold Shelf Life in higher esteem compared to those living within one mile of a grocery store. People living one mile or less from a grocery store are slightly more likely to attend a farmers market or partake in a CSA, which signals a penchant for fresh, high quality produce.

**Table 9:** Produce preferences by respondent group

Rank of Importance	Unconstrained	1 mile or less	Opt to Travel	Over 1 mile	Constrained
<i>First</i>	Freshness	Freshness	Freshness	Freshness	Freshness
<i>Second</i>	Appearance	Appearance	Appearance	Appearance	Appearance
<i>Third</i>	Growing Method	Growing Method	Growing Method	Shelf Life	Shelf Life
<i>Fourth</i>	Seasonality	Shelf Life	Shelf Life	Growing Method	Growing Method^
<i>Fifth</i>	Shelf Life	Seasonality	Seasonality	Seasonality	Seasonality^
Farmers Market or CSA patronage rate	81.9%	79.5%	75.4%	22.2%	76.8%

^Growing Method and Seasonality ranked equally

The “Opt to Travel” group, in contrast to the “Unconstrained” group, preferred to purchase produce with a greater shelf life rather than in-season produce. It should be noted that the “Opt to Travel” group had similar income, household size, vehicle access, and family structure as those in the “Unconstrained” group,<sup>12</sup> yet those that opt to

<sup>12</sup> Average income bracket, car access, household size, and number of children for the “opt to travel” versus “Unconstrained” groups are as follows: \$75,000 to 100,000 income bracket per household for both; 89.8% vs. 94.0% of the sample can access a car whenever they need; 2.6 people per household for both; 0.9 vs. 0.8 children per household.

travel spend less per person per week on groceries (\$49 per person per week versus \$54 per person per week), and grocery shop less frequently (4.4 times per month versus 5.8 times per month) than individuals in the “Unconstrained” group. Thus, respondents in the “Opt to Travel” group appear to have greater preference for value-driven shopping: they buy less fresh produce, hold shelf-life in high regard, shop less often, and stretch their food dollars further within their households compared to those in the “Unconstrained” group.

Respondents in the “Constrained” and “Over 1 mile” groups clearly prioritize extending the shelf life of their purchases, compared to the respondents who live within one mile of a grocery store. This difference may be due to differing preferences between these two groups, but may be a product of their differing food environments. Thus, respondents who live further from a grocery store not only consume more frozen and canned fruits & vegetables relative to those who live near to a grocery store, but also prefer to consume produce with a long shelf-life.

“Constrained” respondents have a surprisingly high farmers market attendance-rate compared with the aggregate of respondents who live over one mile from a grocery store. Since farmers market produce often has significantly longer shelf life than does produce from the grocery store, the “Constrained” respondents may place a high premium on obtaining farmers market produce to meet their need for long-lasting fresh produce.

## Chapter Five: Empirical Analysis and Results

### 5.1 Introduction

In the following chapter, I discuss results from the demand system estimation and MMNL regressions, conduct simulations combining significant results of both methods, and analyze the outcomes. Lastly, I discuss the results of three robustness checks that test the sensitivity of my results to respondent group classifications.

### 5.2 Demand Estimation Results

I estimated the AIDS model separately for the geographic and preference-based constraint analyses. Each model was estimated for six fruit & vegetable categories (fresh vegetables, fresh fruit, canned vegetables, canned fruit, frozen vegetables, and frozen fruit) and all other foods, resulting in a seven-equation system. All elasticities were calculated at the sample medians. Table 10 through Table 13 report expenditure elasticities and Marshallian own- and cross- price elasticities for respondents living within one mile and over one mile from a grocery store, and respondents with “Constrained” as well as “Unconstrained” food geographies.

#### 5.2.1 Geographic Constraint Classification Demand Estimation Results

The results of Table 10 show all positive expenditure elasticities, thus each food category comprises a normal good within this demand system. Consistent with economic theory, nearly all own-price elasticities are negative. Negative own-price elasticity implies that an increase in the price of a good diminishes demand for that good. Canned fruit for the “Over 1 mile” subsample provides the only exception. This anomaly is likely due to the significant infrequency of purchases within this food group. Out of 116 respondents in the “Over 1 mile” sample, 74.1 percent did not purchase any canned fruit in the week or month prior to the survey.

**Table 10:** Estimated expenditure and own-price elasticities for individuals living within 1 mile and over 1 mile of a grocery store

Food Group	1 mile or less		Over 1 mile	
	Food Expenditure	Marshallian own-price	Food Expenditure	Marshallian own-price
(1) Fresh Veg	0.31	-0.41	1.10	-1.02
(2) Fresh Fruit	0.48	-0.76	1.08	-0.98
(3) Canned Veg	1.16	-0.79	1.26	-1.23
(4) Canned Fruit	1.18	-2.12	1.11	4.16
(5) Frozen Veg	0.99	-1.24	1.17	-0.99
(6) Frozen Fruit	1.20	-0.67	1.21	-1.53
(7) Other Foods	1.01	-0.99	0.98	-1.01
Observations	455		116	

In general, the results presented in Table 10 suggest proximity to a grocery store is related to greater consistency of fresh produce consumption. The food expenditure elasticity estimates show “1 mile or less” respondents had

relatively inelastic food expenditure elasticities of 0.31 and 0.48 for fresh vegetables & fruit compared to “Over 1 mile” respondents, where expenditure elasticities were both greater than unity: 1.10 and 1.08, respectively. This result implies that “1 mile or less” individuals treat fresh vegetables & fruit as staple items more so than “Over 1 mile” respondents; demand for fresh vegetables & fruit of the “1 mile or less” group responds minimally to changes in total food expenditure. Conversely, “Over 1 mile” respondents treat “other foods” as staple items more so than “1 mile or less” respondents. “Over 1 mile” respondents had expenditure elasticity less than unity of 0.98 for “other foods” whereas “1 mile or less” respondents had expenditure elasticity above unity of 1.01 for “other foods.”

Results also suggest proximity to a grocery store is related to less price sensitivity for all food groups, and fresh fruit & vegetables in particular. The magnitudes of own-price elasticities are greater among “Over 1 mile” residents for all food groups, excepting frozen vegetables. Thus, “Over 1 mile” residents change consumption of groceries items in response to own-price changes to a greater degree than their “1 mile or less” counterparts. Household income is not a likely driver of this variation in price sensitivity, as both groups had similar average income levels (see Table 4). Among “1 mile or less” respondents, fresh vegetable purchases are the least sensitive to price changes, which implies proximity to a grocery store is correlated with steady fresh vegetable consumption regardless of fresh vegetable price volatility.

Marshallian cross-price elasticities are reported in Table 11. They represent uncompensated demand responses to price and thus denote substitution effects taking into account the consumer’s purchasing power change that results from a price change. Elasticity estimates should be read as the change in quantity demanded of good x by row for the change in price of good y by column. Positive cross-price elasticity indicates net substitutes, while negative cross-price elasticity indicates net complements.

**Table 11:** Estimated Marshallian cross-price elasticities for individuals living within 1 mile and over 1 mile of a grocery store

$e_{x, p(y)}$	1 mile or less							Over 1 mile						
	Fresh Veg	Fresh Fruit	Canned Veg	Canned Fruit	Frozen Veg	Frozen Fruit	Other Foods	Fresh Veg	Fresh Fruit	Canned Veg	Canned Fruit	Frozen Veg	Frozen Fruit	Other Foods
(1) Fresh Veg	<b>-0.50</b>	0.35	0.31	0.30	0.22	0.42	-0.26	<b>-1.02</b>	-0.05	-0.09	-0.04	-0.02	-0.07	0.05
(2) Fresh Fruit	0.02	<b>-0.85</b>	0.12	0.13	-0.01	0.18	-0.23	-0.03	<b>-0.98</b>	-0.07	-0.02	-0.01	-0.05	0.06
(3) Canned Veg	0.06	0.03	<b>-0.94</b>	-0.09	0.08	-0.21	0.08	0.19	0.14	<b>-1.23</b>	0.22	-0.02	-0.17	0.23
(4) Canned Fruit	0.06	0.06	-0.12	<b>-1.23</b>	0.07	0.09	0.10	-0.11	-0.10	-1.07	<b>4.16</b>	-1.07	-2.99	0.17
(5) Frozen Veg	-0.04	-0.10	0.27	0.15	<b>-0.90</b>	-0.35	-0.02	0.00	-0.03	0.04	-0.20	<b>-0.99</b>	-0.12	0.10
(6) Frozen Fruit	0.11	0.05	-0.13	0.05	-0.04	<b>-1.08</b>	0.11	-0.07	-0.05	-0.17	-2.99	-0.12	<b>-1.53</b>	0.20
(7) Other Foods	-0.03	-0.02	0.00	0.00	-0.01	0.01	<b>-0.99</b>	0.05	0.06	0.23	0.17	0.10	0.20	<b>-1.01</b>
Observations	455							116						

The results of Table 11 provide evidence that individuals who live over one mile from a grocery store are more likely to substitute frozen or canned vegetables for fresh vegetables compared to individuals who live within one mile of a grocery store. “Over 1 mile” residents increase consumption of canned vegetables by 0.19 percent for a 1.0 percent increase in the price of fresh vegetables. “1 mile or less” residents substitute non-fresh for fresh vegetables to a lesser degree of 0.06 percent. Also, “1 mile or less” residents treat fresh vegetables and frozen vegetables as complementary goods as indicated by the negative cross-price term of -0.04, while “Over 1 mile” residents exhibit behavior slightly closer to substitution with a positive cross-price elasticity of 0.004 (taken to the thousandth decimal place) for fresh to frozen vegetables. This behavior switches for the fruit categories. “Over 1 mile” residents treat fresh fruit as a complementary good to canned or frozen fruit with cross price elasticities of -0.10 and -0.05, respectively, whereas “1 mile or less” respondents are more likely to substitute fresh for frozen or canned fruits because the cross price terms are positive. For example, for a 1.0 percent price increase of fresh fruit canned fruit demand among “1 mile or less” residents increases by 0.06 percent and frozen fruit increases by 0.05 percent while canned fruit and frozen fruit demand both decrease with a price increase in fresh fruit for the “Over 1 mile” residents. For both groups the relationships between fresh and non-fresh fruits are inelastic, an unsurprising result if we consider that fresh fruit and non-fresh fruits are generally less substitutable for consumption and cooking than are fresh and non-fresh vegetables.

Using this exogenous classification of food access, the above results suggest residents who live within a mile of a grocery store have stable fresh produce consumption and are less price sensitive. People who live over a mile from a grocery store are more likely to substitute frozen or canned vegetables for fresh vegetables, and their consumption of other foods exhibits lower variation than the “1 mile or less” residents. Next, I use an endogenous classification of food access to further test the effect of geography on consumption behavior. Table 12 shows expenditure and own-price elasticities for individuals living with an “Unconstrained” food environment compared to respondents in the “Constrained” group. The last two columns also show expenditure and own-price elasticities for respondents who “Opt to Travel.”

### **5.2.2 Preference-Based Constraint Classification Demand Estimation Results**

The results of Table 12 show all positive expenditure elasticities, thus each food category comprises a normal good within this system. As with the previous analysis, nearly all own-price elasticities comply with economic theory and are negative. Frozen fruit for the “Constrained” and “Opt to Travel” subsamples provide the only exception, which is likely due to the significant infrequency of purchases within this food group. Of these groups, 84.6 of those who “Opt to Travel” and 83.9 percent of those in the “Constrained” group did not purchase any frozen fruit in the week or month prior to the survey. Similarly, the “Constrained” group did not purchase canned fruit 83.9 percent of the time, which may explain the extremely large canned fruit own-price elasticity of -28.07.



**Table 12:** Estimated expenditure and own-price elasticities for Unconstrained, Constrained, and Opt to Travel respondent groups

Food Group	Unconstrained		Constrained		Opt to Travel	
	Food Expenditure	Marshallian own-price	Food Expenditure	Marshallian own-price	Food Expenditure	Marshallian own-price
(1) Fresh Veg	0.33	-0.35	2.40	-2.68	0.73	-0.90
(2) Fresh Fruit	0.70	-0.88	0.94	-0.85	0.69	-1.01
(3) Canned Veg	1.15	-0.90	1.09	-1.44	1.05	-0.83
(4) Canned Fruit	1.16	-1.04	1.43	-28.07	1.01	-2.66
(5) Frozen Veg	1.15	-1.01	1.84	-2.85	0.83	-1.05
(6) Frozen Fruit	1.20	-1.01	1.15	0.77	0.98	0.68
(7) Other Foods	1.02	-0.99	1.03	-1.00	1.03	-0.99
Observations	221		56		260	

Based on these results, it again appears proximity to a grocery store is related to greater consistency of fresh produce consumption. Both the “Unconstrained” and “Opt to Travel” respondents have inelastic expenditure elasticity of demand for fresh produce, compared to those in the “Constrained” group. The “Unconstrained” and “Constrained” groups exhibit demand for canned and frozen fruits & vegetables above unity elastic whereas the “Opt to Travel” group treats frozen fruit & vegetables as relatively more of a staple item. The “Opt to Travel” group may be more likely to frequent big box stores in an effort to garner deals on non-perishable frozen produce. It is unclear whether the endogenous food access classification renders significantly different consumption behavior of non-fresh food items. “Constrained” respondents have lower income elasticity of demand for canned vegetables and frozen fruit, but higher income elasticity of demand for canned fruit and frozen vegetables compared to the “Unconstrained” group. Expenditure elasticity for “other foods” is similar for all groups. Respondents who live within one mile of a grocery store and prefer their status quo have stable demand for fresh produce, but demand for all other food groups is elastic compared to the “Opt to Travel” group.

Results are inconclusive with respect to own-price sensitivity. “Unconstrained” respondents are significantly less price sensitive to fresh vegetables than “Constrained” respondents, but slightly more price sensitive to fresh fruit. Both sets of respondents who live within one mile of a grocery store are less price-sensitive to all non-fresh foods compared to the “Constrained” group, though demand for “other foods” is robustly unit elastic for all groups. While the positive signs on own-price elasticity of frozen fruit (0.77 and 0.68) may be an error caused by non-response bias, it is also plausible that consumers in the “Constrained” and “Opt to Travel” groups purchase frozen fruit as a complement to fresh fruit and thus, do not budget frozen fruit separately from fresh or other food categories. Respondents who “Opt to Travel” are less price-sensitive to fresh vegetables, since the own-price elasticity is -0.90 compared to the more elastic -2.68 of the “Constrained” group who must travel over one mile to reach their grocery store. The opposite is true, however, for fresh fruit. The “Constrained” group is less price-sensitive to fresh fruit with an own-price elasticity of -0.85 whereas the “Opt to Travel” group exhibits fresh fruit

own-price elasticity of -1.01. Thus, respondents who elect to travel are less price- sensitive to all food groups except fresh fruit, compared to those who are forced to travel over one mile.

Added travel time, in addition to added distance to a grocery store, appears to be related to more elastic fresh vegetable consumption. However, in contradiction to my hypothesis, the three groups do not have significantly different demand elasticities for “other foods.”

Marshallian cross-price elasticities for the endogenous “food desert” analysis are reported in Table 13. Table 14 provides cross-price elasticities for respondents who “Opt to Travel.” Recall, a positive cross-price elasticity indicates net substitutes, while a negative cross-price elasticity indicates net complements.

**Table 13:** Estimated Marshallian cross-price elasticities for Unconstrained and Constrained respondent groups

$e_{x, p(y)}$	Unconstrained							Constrained						
	Fresh Veg	Fresh Fruit	Canned Veg	Canned Fruit	Frozen Veg	Frozen Fruit	Other Foods	Fresh Veg	Fresh Fruit	Canned Veg	Canned Fruit	Frozen Veg	Frozen Fruit	Other Foods
<b>Food Group</b>														
(1) Fresh Veg	<b>-0.35</b>	0.31	0.36	0.33	0.28	0.49	-0.29	<b>-2.68</b>	-1.82	-2.49	-2.04	-1.69	-2.24	-3.08
(2) Fresh Fruit	-0.12	<b>-0.88</b>	0.05	0.01	0.05	0.11	-0.14	-0.02	<b>-0.85</b>	0.07	0.00	-0.08	0.03	0.04
(3) Canned Veg	0.09	0.03	<b>-0.90</b>	0.13	-0.10	-0.32	0.08	0.14	0.11	<b>-1.44</b>	-0.47	0.64	0.22	-0.06
(4) Canned Fruit	0.07	-0.05	0.37	<b>-1.04</b>	-0.02	-0.42	0.12	0.71	0.80	6.93	<b>-28.07</b>	4.15	15.06	-0.77
(5) Frozen Veg	0.00	0.03	-0.10	-0.01	<b>-1.01</b>	0.00	0.10	1.76	1.76	-5.58	2.37	<b>-2.85</b>	1.78	-1.38
(6) Frozen Fruit	0.13	0.04	-0.15	-0.06	0.01	<b>-1.01</b>	0.12	0.13	0.06	0.34	-1.66	-0.33	<b>0.77</b>	-0.11
(7) Other Foods	-0.04	-0.02	-0.01	-0.01	-0.01	0.00	<b>-0.99</b>	0.05	0.03	0.02	0.03	0.03	0.02	<b>-1.00</b>
Observations	221							56						

**Table 14:** Estimated Marshallian cross-price elasticities for Opt to Travel group

$e_{x, p(y)}$	Opt to Travel						
	Fresh Veg	Fresh Fruit	Canned Veg	Canned Fruit	Frozen Veg	Frozen Fruit	Other Foods
<b>Food Group</b>							
(1) Fresh Veg	<b>-0.90</b>	0.30	0.37	0.33	0.29	0.37	0.58
(2) Fresh Fruit	0.09	<b>-1.01</b>	0.27	0.23	0.16	0.24	0.58
(3) Canned Veg	0.15	0.13	<b>-0.83</b>	-0.03	0.12	-0.42	-0.11
(4) Canned Fruit	0.00	0.04	-0.11	<b>-2.66</b>	1.10	0.67	-0.04
(5) Frozen Veg	-0.27	-0.24	0.32	0.82	<b>-1.05</b>	-0.55	0.13
(6) Frozen Fruit	0.71	0.28	-2.30	1.01	-1.20	<b>0.68</b>	-0.19
(7) Other Foods	0.06	0.05	0.04	0.04	0.04	0.04	<b>-0.99</b>
Observations	260						

As I expect, the results of Table 13 demonstrate that the “Constrained” group is more likely to substitute fresh for non-fresh produce compared to the “Unconstrained” group. Respondents in the “Constrained” category exhibit a higher elasticity of substitution for both fresh-to-canned vegetables of 0.14 and fresh-to-frozen vegetables of 1.76 compared to “Unconstrained” category respondents. Similarly, the “Constrained” respondents have greater elasticity of substitution for fresh-to-canned fruit of 0.8 and fresh-to-frozen fruit of 0.6 compared to respondents in the “Unconstrained” category. Both sets of respondents treat “other foods” as complementary to fresh vegetables, as the cross-price elasticities are both negative. However, the “Constrained” group is more likely to

substitute fresh fruit for “other foods” with a positive cross-price elasticity of 0.04 whereas the “Unconstrained” group treats “other foods” and fresh fruit as complementary goods with a negative cross-price elasticity of -0.14. This result provides evidence that the “Constrained” group is more sensitive to perishability than the “Unconstrained” group.

Further Table 13 illustrates that constrained food geography intensifies the propensity to substitute non-fresh for fresh produce. The “Constrained” group substitutes fresh vegetables (fruit) for canned or frozen vegetables (fruit) to a greater degree than the “Over 1 mile” group.

Comparing behavior between those who are forced to travel (the “Constrained” group) and those who elect to travel (the “Opt to Travel” group) in Table 14 provides inconclusive results. The “Constrained” group more readily substitutes frozen for fresh vegetables with a cross price elasticity of 1.76 compared to -0.27 for the “Opt to Travel” group. Yet, both groups substitute fresh for canned vegetables to the same degree, with an elasticity of 0.14 and 0.15, respectively. Similarly, the “Constrained” group substitutes fresh for canned fruit to a greater degree than those who “Opt to Travel” but substitutes fresh for frozen fruit to a lesser degree. It should be noted that the “Constrained” cross- price elasticity of frozen-to-fresh vegetables of 1.76 indicates highly elastic substitution above unity. Thus, fresh to frozen vegetable substitution appears to be the most susceptible to change due to geographic constraint.

The highly elastic cross-price elasticity terms on fresh vegetables for the “Constrained” group indicates these respondents are likely to compromise their purchase of fresh vegetables more than any other food category if they experience a budget squeeze. They essentially treat fresh vegetables as a residual category. If the price increases on any other food category, they compensate by purchasing fewer fresh vegetables. Some of this result may be driven by aversion to vegetables. The highly elastic income- elasticity term of frozen vegetables in Table 12 would support this plausible hypothesis. However, the group has signaled that they are dissatisfied with some aspect of their grocery shopping experience, which may be related to produce shelf life. A typical basket of fresh vegetables available in central Illinois grocery stores tends to perish faster than a typical basket of fresh fruit. Lettuce, peppers, and cucumbers, for example, have significantly shorter shelf life than do apples, bananas, and oranges. Recall the “Constrained” group regarded shelf life as an important aspect of produce quality (see Table 9). The “Constrained” group’s highly elastic vegetable consumption may be a product of poor quality or highly perishable vegetables. Survey respondents were asked: “Are there food products you would like to purchase regularly that are not available?” and among the “Constrained” group, 26.3 percent responded to this question with fresh fruit or fresh vegetables. Comparatively, only 9.1 percent of the total sample responded with fresh fruit or fresh vegetables. These summary statistics are evocative of significant influences that inhibit fresh produce consumption of the “Constrained” group, which may explain their elastic fresh produce demand.

Respondents who “Opt to Travel” substitute fresh for canned vegetables to a greater degree than the “Unconstrained” respondent group with a cross-price elasticity of 0.15 compared to 0.09, and treat fresh and frozen vegetables as complements with a negative cross-price elasticity of -0.27, unlike the “Unconstrained” group, who treat the two food categories independently. The “Opt to Travel” respondents substitute canned and frozen fruit for fresh fruit to a greater degree than does the “Unconstrained” group. For example, “Opt to Travel” respondents increase consumption of canned fruit by 0.04 percent for a one percent increase in price of fresh fruit, and increase consumption of frozen fruit by 0.28 percent for a one percent increase in price of fresh fruit. Whereas respondents in the “Unconstrained” group actually decrease canned fruit consumption by 0.05 percent and increase frozen fruit to a lesser degree of 0.04 percent for a one percent increase in the price of fresh fruit. If we consider the effect of a price decrease in “other foods,” people who “Opt to Travel” will decrease their demand of fresh fruits & vegetables by 0.58 percent, whereas respondents in the “Unconstrained” group will respond by increasing demand for fresh vegetables by 0.29 percent and fresh fruit by 0.14 percent.

### **5.2.3 Summary of Demand Estimation Results**

Respondents who live near to a grocery store exhibit greater stability of fresh produce consumption, compared to those who live over one mile. Proximate respondents also are less sensitive to price volatility of fresh vegetables, as well as non-fresh fruit & vegetable categories. Constrained food geography may intensify the propensity to substitute non-fresh for fresh produce. Such is the case when comparing the significantly different respondent groups of “Constrained” versus “Unconstrained” respondents. Proximity to a grocery store may have a positive effect on consumption behavior, even if an individual does not patronize the store nearby their residence. The “Constrained” respondents substitute frozen for fresh vegetables to a significantly greater degree than do respondents who also travel far to get to their grocery store, but do so willingly, e.g. the “Opt to Travel” respondents. Proximity to a grocery store is not the most likely cause of the divergent consumption behavior because the “Opt to Travel” group are traveling similar distances and shopping at similar stores as the “Constrained” group. Table 4 illustrates both groups frequently patronize supercenters. Also, the average stated travel time for the “Opt to Travel” group and the “Constrained” groups are 12.1 minutes and 12.3 minutes, respectively.

This result has two important implications. First, the products people purchase at grocery stores can vary significantly, thus simply adding grocery stores to underserved areas will not necessarily induce local residents to purchase more fresh produce, even if they do patronize the store. In other words, patronage of a grocery store does not necessitate fresh produce consumption. Second, proximity to a grocery store does not necessitate patronage of that grocery store. This provides evidence to discredit Hypothesis 3: additional travel time to a grocery store is not uniquely responsible for fresh produce consumption.

In summary, the geographic and behavioral constraint definitions reveal proximity to a grocery store is positively related to stable consumption of fresh produce, and lower variation of fresh produce consumption in response to price changes. These results affirm Hypothesis 1, described in Section 3.2.1. The income elasticity results reflect the findings of Weatherspoon et al. (2013) who found food residents had a high income- elasticity of demand for fresh fruit. The cross-price elasticity results provide some evidence that proximity to a grocery store decreases the likelihood of substituting fresh for non-fresh vegetables, even if people do not shop at the store nearest their house. Further, constrained food geography exacerbates substitution behavior between fresh and non-fresh produce. This result affirms Hypothesis 2. Previous literature supports this finding. Ver Ploeg et al. (2009), Rose and Richards (2003), and Wrigley et al. (2002) each found that increased distance from a grocery store or infrequent grocery shopping were correlated with lower consumption of fresh produce. The following discussion explores whether the observed consumption behavior may be caused by food geography or consumer preferences.

### 5.3 Choice Experiment Results

The choice experiments quantify consumer preferences for grocery store attributes. I compare these results with the previous demand system results to test whether “Over 1 mile” or “Constrained” respondents are constrained by their food geography. Following the previous methodology, I use both a preference-based and a geographic constraint classification to test for preferential differences.

I estimated the MMNL model with the full sample prior to splitting the sample into groups. Next, I estimate three sets of MMNL models to compare with the demand systems. The first set compares preferences between “Over 1 mile” and “1 mile or less” respondents following the geographic constraint classification. The second set includes the subset of “Over 1 mile” and “1 mile or less” respondents who were less likely than average to choose their status quo shopping scenario – the “Constrained” respondents, and more likely than average to choose their status quo shopping scenario – the “Unconstrained” respondents, respectively following the preference-based constraint classification. The third set includes the “Constrained” and “Opt to Travel” respondents. For all three cases, I report parameter estimates of the MMNL model as well as marginal willingness-to-pay calculations.

The models presented in this chapter represent the best-fit specifications I achieved after several iterations. I estimated a main effects-only model for all three cases. Goodness of fit was based on comparing the AIC and BIC values across models, as well as the robustness of each parameter estimate.<sup>13</sup> Parameter estimate results for these main-effects models are provided in Appendix E.

---

<sup>13</sup> AIC (Akaike’s Information Criterion) and BIC (Bayesian Information Criterion) are common statistical measures used in probability-based regressions to assess the fit of a model based on measures of information. When comparing two models, the

### 5.3.1 Geographic Constraint Classification Choice Experiment Results

To compare grocery store preferences between “Over 1 mile” and “1 mile or less” residents, I multiply each main effect parameter in equation (12) by a dummy variable, denoted by *FD* for “food desert”, equal to 1 if the respondent lives over one mile from a grocery store, and equal to 0 if the respondent lives less than one mile from a grocery store.<sup>14</sup>

Results of the geographic constraint MMNL regressions are presented in Table 15. The first two columns show regression results for the full sample of respondents. The second two columns incorporate the *FD* dummy variable to highlight differences between “Over 1 mile” and “1 mile or less” respondents. The column marked “SD Pval” provides the p-value associated with the standard deviation of each parameter. Each of the main effect parameters has significant standard deviation around their mean, which implies individual heterogeneity is significant for all attributes. This result justifies employing the MMNL model as opposed to the CL model because the MMNL model yields more conservative estimates of marginal WTP compared to the CL when individual heterogeneity is significant.

Considering the full population results, all main effect variables are significant at one percent, excepting the *Grocery Store* and *AQ Produce* variables. The ASC term captures attributes affiliated with the status quo choice that I did not explicitly account for in the experimental design. ASC is negative and significant meaning respondents had disutility from choosing either new store option A or B. Thus, respondents preferred to choose their current store rather than the new hypothetical stores. This result follows Tang et al. (2001) who found knowledge of store layout was a significant source of utility of grocery shopping. The negative sign on *Weekly Grocery Bill* and *Travel Time* meet my a priori expectations that people will have disutility from spending more money on groceries or taking extra time to travel to a grocery store, ceteris paribus. *Specialty Store* and *Grocery Store* are both positive, which suggests respondents are more likely to choose to shop at either a specialty store or grocery store instead of a supercenter, though the *Grocery Store* attribute is statistically insignificant. The respondents’ disutility from supercenters mimics the findings of Brown (1978) who found travel distance had a stronger effect than store size in determining store patronage. In this study, grocery stores were closer to respondent homes on average than supercenters (see Table 3). This tendency to choose smaller stores also follows Brown’s (1978) finding that grocery

---

smaller the value of the AIC or BIC statistic, the better the fit of the model. BIC identifies the model that is more likely to have generated the observed data. The AIC offers a relative estimate of improved fit versus deviance (Williams, Richard 2011).

<sup>14</sup> The four interaction terms from equation (12)  $\beta_7 X_{iBill} * Y_i CurrentTravel$ ;  $\beta_8 X_{iBill} * Y_i Distance$ ;  $\beta_9 X_{iHQProduce} * Y_i Income$ ; and  $\beta_{10} X_{iAQProduce} * Y_i Income$

become “triple” interaction terms when multiplied by the *FD* dummy. After running the MMNL including these triple interaction terms, each triple interaction parameter estimate was insignificant. This suggests the interaction effects of distance and travel time on marginal WTP, and income on produce quality are not significantly different for the “1 mile or less” group compared to the “Over 1 mile” group. Consequently, I dropped the triple interaction terms. Model fit improved without the triple interaction terms, since the AIC and BIC measures were consistently lowest in the MMNL model without triple interaction effects.

store size attracts patronage up to a point, after which increasing size deters grocery shopping. *HQ Produce* is positive and significant, which indicates as I expect, that people garner higher utility from high quality produce compared to low quality produce. Availability of average quality produce instead of low quality produce does not significantly affect respondent store selection.

Interaction terms *Distance x Grocery Bill* and *CurrentTravel x Grocery Bill* are negative and significant. This implies that respondents who live further from a grocery store, or have a longer travel time will have a lower probability of choosing a store with a high grocery bill. The added distance or travel time, itself, adds a cost premium to the grocery bill, so it is not surprising that respondents faced with this additional “cost” to grocery shop would be averse to expensive groceries. The interactions of *Income x HQ Produce* as well as *Income x AQ Produce* are both positive and significant, which indicates higher income is correlated with greater likelihood of choosing a store with better quality produce. This finding reflects that of Wilde, McNamara, and Ranney (1999 and 2000) who found low-income households were more likely to consume non-fruit & vegetable products, even if fruits & vegetables were subsidized.

The results of each main effect are robust between the “MMNL-Full Sample” model and the “MMNL – Geographic Classification model,” i.e. the MMNL with the *FD* dummy. If the main effect variables interacted with the *FD* dummy are statistically significant, “Over 1 mile” respondents have significantly different utility compared to the “1 mile or less” group for that main effect. The change in utility for a change in *HQ Produce* alone is significantly different for “Over 1 mile” respondents compared to respondents who live within one mile of a grocery store. “Over 1 mile” respondents experience more utility from high quality produce than do the “1 mile or less” respondents. For example, shopping at a store with high instead of low quality produce increases the probability of choosing that store by a factor of  $2.165 + 1.266 = 3.431$  among “Over 1 mile” respondents. While it may seem counterintuitive that the group with less consistent fresh produce consumption would gain higher utility from high quality produce, this result supports an argument for exogenous food geography constraints. For example, it is possible that the average produce quality currently observed by the “Over 1 mile” group is subpar to their expectations on average, thus they garner higher utility from an increase in produce quality relative to the “1 mile or less” group, who may be currently very pleased with their produce quality.

A noteworthy result of Table 15 is the insignificance of *Travel Time\_FD*, which implies that respondents who live far from a grocery store do not have significantly greater disutility from added travel time than do those who live near to grocery stores. In fact, the *Travel Time\_FD* parameter estimate is positive, suggesting “Over 1 mile” respondents have less disutility from added travel time though the difference is not statistically significant. Despite current policy concern for the adverse effect of added travel costs on fresh produce consumption, these results suggest travel time itself may have little effect on produce consumption.

Respondents who live further from a grocery store garner more disutility from an increased grocery bill. Two major results from the MMNL regression support this conclusion. First, though the *Weekly Grocery Bill\_FD* term is insignificant, the point estimate for the full effect of the grocery bill attribute must be calculated through a linear combination of grocery bill estimators. I used the Stata command `lincom` to test whether the linear combination of grocery bill estimators was statistically different from zero. Results of the linear combination of estimators significance test are provided in Appendix F. These test results show the weekly grocery bill effect is negative and significant for the “Over 1 Mile” group, which implies they have slightly more disutility from a higher grocery bill than do the “1 mile or less” group. In other words, the “Over 1 mile” group is less likely to choose a store option with a higher grocery bill than the “1 mile or less” group, all else equal.

Second, the negative and significant *Distance x Grocery Bill* interaction term signifies that the “Over 1 mile” respondent group is more price-sensitive than are those who live close to a grocery store. Greater disutility from a higher grocery bill is consistent with my findings from the demand estimation in Section 5.2 where “Over 1 mile” residents were more price-sensitive for groceries relative to the “1 mile or less” group. Price-sensitivity for the “Over 1 mile” group is not likely attributable to income constraints, however, since on average the “Over 1 mile” group has higher income than do the “1 mile or less” group (see Table 4).



**Table 15: MMNL results with geographic constraint classification**

Variable	MMNL - Full Sample			MMNL - Geographic Classification		
	Coef (SE)	SD	Pval <sup>a</sup>	Coef (SE)	SD	Pval <sup>a</sup>
<b>Main Effects</b>						
ASC	-2.297 *** 0.107			-2.344 *** 0.123		
Weekly Grocery Bill	-0.027 *** 0.007			-0.031 *** 0.008		
Travel Time	-0.129 *** 0.010	0.000		-0.141 *** 0.012	0.000	
Specialty Store	0.487 *** 0.140	0.000		0.409 ** 0.166	0.000	
Grocery Store	0.052 0.108	0.000		-0.027 0.120	0.000	
HQ Produce	2.284 *** 0.445	0.000		2.165 *** 0.465	0.000	
AQ Produce	-0.554 0.387	0.000		-0.748 * 0.402	0.093	
<b>Main Effect 'Food Desert' Dummies</b>						
ASC_FD				-0.509 0.354	0.000	
Weekly Grocery Bill_FD				0.003 0.017	0.069	
Travel Time_FD				0.010 0.025	0.002	
Specialty Store_FD				0.171 0.395	0.005	
Grocery Store_FD				0.266 0.316	0.023	
HQ Produce_FD				1.266 *** 0.460	0.003	
AQ Produce_FD				0.183 0.380	0.263	
<b>Interaction Effects</b>						
Distance x Grocery Bill	-0.007 *** 0.002	0.002		-0.010 *** 0.004	0.289	
CurrentTravel x Grocery Bill	-0.002 ** 0.001	0.000		-0.001 0.001	0.000	
Income x HQ Produce	0.261 *** 0.083	0.005		0.269 *** 0.089	0.137	
Income x AQ Produce	0.267 *** 0.072	0.474		0.302 *** 0.077	0.032	
<b>Observations</b>						
Log Likelihood	9,609 -2279.30			9,609 -2250.91		
Prob > Chi2	0.000			0.000		
AIC	4580.59			4537.82		
BIC	4659.47			4666.89		

\*\*\* significant at 1%; \*\* significant at 5%; \* significant at 10%

<sup>a</sup>Significance of standard deviations of individual heterogeneity

Coefficient estimates for remaining attributes are better interpreted through a marginal WTP measure. Marginal WTP values for the full population and *FD* dummy MMNL models are shown in Table 16. Unlike the coefficient estimates presented in Table 15, the marginal WTP values depend on respondent's disutility from the weekly grocery bill. Thus, the marginal WTP value is related to, but should be interpreted differently from a utility measure. Unit values for the respondent characteristic variables of travel time, distance, and income are based on the sample average because I am interested in the marginal WTP for the average person in each group: full sample, "Over 1 mile" respondents, and "1 mile or less" respondents. In other words, I calculate the marginal WTP per grocery shopping trip for the average respondent in these groups.

**Table 16: Marginal Willingness-to-Pay – Geographic constraint classification**

	Full Sample	1 mile or less	Over 1 mile	Significant Difference?*
Travel Time (\$/min. travel)	-\$2.50	-\$3.08	-\$1.51	No
Specialty Store (\$/specialty vs supercenter)	\$9.44	\$8.95	\$6.73	No
Grocery Store (\$/grocery vs supercenter)	\$1.00	-\$0.58	\$2.78	No
HQ Produce (\$/HQ vs LQ produce)	\$71.62	\$78.95	\$58.03	Yes
AQ Produce (\$/AQ vs LQ produce)	\$17.26	\$19.02	\$13.91	No
<b>Unit Values</b>				
Average Travel Time (min.)	9.4	8.6	12.3	
Average Distance (mi.)	1.2	0.5	4.3	
Average Income Bracket <sup>^</sup>	5.4	5.4	5.9	

\*Significant difference between sample groups of utility for the attribute, not marginal WTP for the attribute

<sup>^</sup>Income bracket of 4 denotes annual income between \$50-75K; 5 denotes \$75-100K; 6 denotes \$100-125K

The marginal WTP for added travel time is smaller for the “Over 1 mile” respondents than for the “1 mile or less” respondents. Individuals living over one mile from a grocery need to be compensated \$1.51 to travel an extra minute as opposed to “1 mile or less” individuals who need to be compensated \$3.08. This result supports an endogenous sorting explanation whereby people who choose to live closer to a grocery store may do so because they value short trips to the grocery store, whereas individuals who live further away find other locational amenities, such as acreage, more desirable than proximity to a grocery store.

“1 mile or less” respondents are willing to pay more, \$8.95, to shop at a specialty store instead of a supercenter compared to “Over 1 mile” respondents who are willing to pay a slightly smaller amount of \$6.73. Conversely, the “1 mile or less” respondents require compensation of 58¢ to shop at a grocery store instead of a supercenter. This result may be explained by the fact that one of the most popular grocery stores in Champaign-Urbana is indeed a supercenter model, a regional retailer called Meijer.<sup>15</sup> Also, it should be noted that my choice experiment described “supercenter” as a retailer characterized by their size, the number of non-food items available, and the number of extra services such as pharmacy or florist that are available. I did not provide any inferences on product quality or ownership structure in my definition of store type. The respondent’s preference for supercenters mimics the findings of Tang et al. (2001) who found that a larger product assortment in a store including breadth and depth resulted in a greater likelihood that a consumer would choose that store because large assortment facilitates “one-stop shopping” and obviates the need to make separate trips to other stores. In other words, it reduces the fixed costs associated with grocery shopping.

<sup>15</sup> Of all grocery stores in Champaign County, the highest share of respondents chose Meijer as their primary grocery store (30.9%), followed by Schnucks (30.6%) and County market (13.9%).

“Over 1 mile” respondents are willing to pay more to shop at a grocery store compared to “1 mile or less” respondents despite the fact that “Over 1 mile” respondents were more likely than “1 mile or less” respondents to have shopped at a supercenter in the month prior to the survey. Thus, preferences for store type appear at odds with behavior for the “Over 1 mile” group, which lends some evidence to the existence of an exogenous food constraint.

Another striking result from Table 16 is the significant difference in willingness-to-pay for high quality produce over average quality produce across all sample groups. Marginal WTP for high quality produce is quite high. This result is driven by the large, positive coefficient on *HQ Produce* in each MMNL model. “1 mile or less” respondents exhibited higher marginal WTP for both average and high quality produce compared to the “Over 1 mile” respondents. Yet, “Over 1 mile” respondents garner significantly greater utility from higher produce quality than do the “1 mile or less” respondents, since the *HQ Produce\_FD* parameter estimate is both positive and significant. Preferences for produce of the “Over 1 mile” group, therefore, appear to be inconsistent with the results of the demand system in Section 5.2 in which the “Over 1 mile” respondents displayed less consistent fresh fruit & vegetable consumption, which again provides evidence that exogenous treatment effects of the food environment are partially attributable for the divergent consumption behavior between the “Over 1 mile” and “1 mile or less” groups.

Neither travel time nor income appear to constrain the “Over 1 mile” group, however, their relatively inconsistent fruit & vegetable consumption and propensity to substitute frozen or canned vegetables for fresh vegetables, are somewhat anomalous with their preferences for high quality produce. The “Over 1 mile” group garners higher utility from high quality produce than does the “1 mile or less” group since the *HQ Produce\_FD* parameter estimate is positive and significant. Nonetheless, they are less willing-to-pay for quality in absolute terms compared to the “1 mile or less” group. Consequently, it would be difficult to conclude based on these results that produce quality would drive these respondents from shopping at a supercenter to shopping at a grocery store if grocery stores were more accessible. In other words, locational sorting and treatment effects both appear to drive consumption behavior. The following analysis attempts to parse such effects by accounting for respondents’ perception of their food environment.

### 5.3.2 Preference-Based Constraint Classification Choice Experiment Results

In the following analysis, I consider a subset of the previous analysis and compare respondents who live within one mile of a grocery store who indicated satisfaction with their current food geography by selecting the status quo option more often than average, the “Unconstrained” group; and respondents who live over one mile of a grocery store who indicated dissatisfaction with the current food geography by selecting the status quo option less often than average, the “Constrained” group.

To compare grocery store preferences between “Constrained” and “Unconstrained” respondents, I multiply each main effect term of equation (12) by a dummy variable, denoted by  $CD$ , equal to 1 if the respondent is in the “Constrained” group, and equal to 0 if the respondent is in the “Unconstrained” group.<sup>16</sup>

Results of the preference-based MMNL regressions are presented in Table 17. Regression results for a pooled sample of the “Unconstrained” and “Constrained” group is uninformative, thus I show only regression results for the MMNL incorporating the  $CD$  dummy variable to highlight where preferences diverge for the two sample groups. The sign, significance, and magnitude of each main effect coefficient are similar to that of Table 16 thus results for the “1 mile or less” respondents, which here correspond to the “Unconstrained” respondent group, follow the discussion in Section 5.3.1 with a few exceptions. The *Specialty Store* attribute is insignificant for both groups, which implies neither the “Constrained” or “Unconstrained” respondents garner significantly different utility from shopping at a grocery or specialty store instead of a supercenter store. Results of Table 17 suggest, as before, that greater distance from a grocery store is correlated with greater aversion to a higher grocery bill.  $Distance \times Grocery\ Bill$  is negative and significant. The negative value implies that respondents who live further away from their nearest grocery store are more sensitive to a higher grocery bill. This implies a continuous relationship between distance and bill sensitivity: respondents who live 0.75 miles from a grocery store are more bill sensitive than those who live 0.25 miles. Additionally, the “Constrained” respondents garner more disutility from a higher grocery bill than the “Unconstrained” respondents. The  $Weekly\ Grocery\ Bill\_CD$  coefficient in Table 17 is insignificant, however, after testing for significance using the linear combination of grocery bill estimators, the full effect of the grocery bill is negative and significant for the “Constrained” group. The grocery bill linear combination calculation and significance test are provided in Appendix F.

As with the geographic constraint classification, only the  $HQ\ Produce\_CD$  term is statistically significant among the main effect  $CD$  dummy interaction terms. This implies that only the change in utility from high quality produce is

---

<sup>16</sup> As with the previous analysis, each triple interaction term was insignificant, implying the effects of distance and travel time on marginal WTP, and income on produce quality are not significantly different for respondents in the “Constrained” group compared to respondents in the “Unconstrained” group. Consequently, I dropped the triple interaction terms for simplicity to calculate the marginal WTP for grocery store attributes. Again, the simpler model appeared to have the best fit based on the low AIC and BIC values.

statistically different for respondents in the “Constrained” group compared to respondents in the “Unconstrained” group. Also, the insignificance of *Travel Time\_CD*, again, implies travel time alone may have little effect on produce consumption. The *ASC\_CD* term is positive and significant which signifies the “Constrained” respondents gain more utility from selecting an alternative, hypothetical grocery store option instead of their status quo option compared to the “Unconstrained” group. This result is in sharp contrast to the highly negative *ASC* term, which shows the “Unconstrained” respondents garner substantial disutility from choosing a hypothetical store as opposed to their status quo store.

**Table 17:** MMNL results with preference-based constraint classification

Variable	MMNL - Preference Classification		
	Coef (SE)	SD	Pval <sup>a</sup>
<b>Main Effects</b>			
ASC	-4.150 *** 0.285		
Weekly Grocery Bill	-0.027 *** 0.010		
Travel Time	-0.136 *** 0.021		0.001
Specialty Store	0.060 0.248		0.002
Grocery Store	-0.170 0.166		0.689
HQ Produce	3.139 *** 0.902		0.080
AQ Produce	0.244 0.873		0.874
<b>Main Effect 'Constrained' Dummies</b>			
ASC_CD	3.582 *** 0.415		0.304
Weekly Grocery Bill_CD	-0.014 0.023		0.011
Travel Time_CD	0.046 0.030		0.874
Specialty Store_CD	0.537 0.503		0.043
Grocery Store_CD	0.240 0.341		0.046
HQ Produce_CD	1.803 ** 0.817		0.003
AQ Produce_CD	0.596 0.540		0.097
<b>Interaction Effects</b>			
Distance x Grocery Bill	-0.010 ** 0.005		0.235
CurrentTravel x Grocery Bill	0.000 0.001		0.696
Income x HQ Produce	0.018 0.147		0.862
Income x AQ Produce	0.008 0.137		0.633
<b>Observations</b>			
Log Likelihood	4,614		
Prob > Chi2	-696.86		
AIC	0.007		
BIC	1429.72		
	1545.58		

\*\*\* significant at 1%; \*\* significant at 5%, \* significant at 10%

<sup>a</sup>Significance of standard deviations of individual heterogeneity

I continue discussion of preferential differences between respondents in the “Unconstrained” and “Constrained” groups using their marginal WTP for grocery store attributes. Marginal WTP values for the preference-based constraint MMNL model are shown in Table 18. As with the previous analysis, unit values for the respondent characteristic variables of travel time, distance, and income are based on the average for each sample group.

**Table 18:** Marginal Willingness to Pay – Preference-based constraint classification

	Unconstrained	Constrained	Significant Difference?*
Travel Time (\$/min. travel)	-\$4.21	-\$1.07	No
Specialty Store (\$/specialty vs supercenter)	\$1.86	\$7.13	No
Grocery Store (\$/grocery vs supercenter)	-\$5.27	\$0.84	No
HQ Produce (\$/HQ vs LQ produce)	\$100.64	\$60.29	Yes
AQ Produce (\$/AQ vs LQ produce)	\$8.91	\$10.57	No
<b>Unit Values</b>			
Average Travel Time (min.)	7.4	12.3	
Average Distance (mi.)	0.5	4.1	
Average Income Bracket <sup>^</sup>	5.6	5.9	

\*Significant difference between sample groups of utility for the attribute, not marginal WTP for the attribute

<sup>^</sup>Income bracket of 4 denotes annual income between \$50-75K; 5 denotes \$75-100K; 6 denotes \$100-125K

The respondents in the “Unconstrained” group appear to place substantially higher premium on their time than do respondents in the “Constrained” group, requiring a compensation of \$4.21 compared to \$1.07 for the “Constrained” group. This result mimics that of the geographic constraint analysis, and supports an endogenous sorting explanation whereby people who choose to live closer to a grocery store may do so because they value short trips to the grocery store.

Using the preference-based constraint classification, the differences in marginal WTP for each store type have become wider across respondent groups. The “Over 1 mile” group WTP to grocery shop at a specialty store instead of a supercenter was formerly over \$2.22 lower than that of the “1 mile or less” group (\$6.73 - \$8.95). Now, the “Constrained” group is willing to pay over \$5 more than the “Unconstrained” group to shop at a specialty store over a supercenter (\$7.13 - \$1.86). Willingness-to-pay to shop at a grocery store experienced a similar change. The “Constrained” group has the highest WTP for the *Specialty Store* attribute of all respondent groups exhibited thus far. This result is particularly surprising given the “Constrained” group had the lowest incidence of shopping at specialty stores and the highest incidence of shopping at supercenters compared to other respondent groups, as presented in Table 4.

Similar to the previous geographic constraint analysis, respondents perceive the quality differential between low and high quality to be significantly different from the quality differential between low and average quality. The “Unconstrained” group perceives high quality produce to be over 11 times more valuable than average quality produce based on their marginal WTP of \$8.91 for average and \$100.64 for high quality. The “Constrained” group has a lower marginal WTP for high quality produce compared to the “Unconstrained” group, yet they have a higher

marginal WTP for average quality produce. Further, the “Constrained” group perceives the quality premium between high quality and average quality produce to be approximately six times as much, a substantially lower differential than that of the “Unconstrained” group. Their marginal WTP for average over low quality is \$10.57, and their marginal WTP for high over low quality is approximately six times higher at \$60.29. The “Constrained” group not only values average quality produce more than the “Unconstrained” group, but values average quality produce relative to high quality produce to a higher degree than the “Unconstrained” group.

Recall from Table 12 and Table 13 the “Constrained” group had highly elastic demand for fresh and frozen vegetables. Their tendency to treat vegetables as a residual category may indeed be driven by legitimate displeasure from vegetables as opposed to some exogenous constraint. However, a supplementary analysis on the choice experiment results revealed the “Constrained” group had a higher propensity to select store options with “high quality” produce relative to other respondents. Appendix D provides matrices categorizing respondent choice experiment selections according to their relative produce quality. Of the “Constrained” group, 54.3 percent chose an alternative with high quality produce, compared to 22.9 percent and 13.8 percent for the full sample and “Unconstrained” groups, respectively. The “Constrained” group clearly preferred store options with high quality produce. Some of the apparent differences in quality value perception between the “Unconstrained” and “Constrained” groups may be explained by unobserved endogeneity. The “Unconstrained” group may have better access to high quality products where the “Constrained” group may have limited or no access to products they consider to be high quality. Consequently, their limited consumption of very high quality produce diminishes the value they can hypothetically place on such produce because they may be less familiar with high quality produce relative to their “Unconstrained” food geography counterparts.

Based on these results, the “Constrained” group appears to be dissatisfied with the amenities of their current grocery retailers. Recall that 89.3 percent of the “Constrained” group shopped at a supercenter in the month prior to the survey, a proportion well above the sample average of 60.3 percent (see Table 4). Yet, this group would actually prefer to shop at a specialty store instead of a supercenter. Also, the “Constrained” respondents were least likely to have shopped at a specialty store in the month prior to the survey compared to other sample groups, thus it appears this group is dissatisfied with their lack of access to specialty stores. Travel time, on the other hand, does not appear to be the constraining factor. The “Constrained” group is no more sensitive to added travel time than the “Unconstrained” group, thus their dissatisfaction may originate from some other grocery store attribute such as store ambience, layout, or product offerings. Income clearly is not driving the “Constrained” group to shop at supercenters, since they had the largest mean and medium income of all other groups (see Table 4). Based on their WTP for specialty and grocery store types, it appears the “Constrained” group indeed would prefer to shop at a store type different from a supercenter, but geographic constraints prevent them from doing so. To what extent increased specialty store patronage would change their fresh produce consumption compared to supercenter



patronage is an open question. Yet, if we believe the findings of Blitstein, Snider, and Evans (2011), as well as the correlations suggested by my own data, the more satisfied a consumer is with their produce retailer, the more fresh fruits & vegetables they will consume. The rank of respondent satisfaction with their current grocery stores – as indicated by the number of times they selected the status quo option – follows “Unconstrained”, “1 mile or less”, “Over 1 mile”, and “Constrained” (see Table 3). This ordering is identical when considering the grocery budget shares devoted to fresh produce (see Table 7).

In the following Section 5.3.4 I perform a robustness check on preferences of the “Constrained” group to test which grocery store attribute may have the greatest influence on fresh produce consumption.

### **5.3.3 Summary of Choice Experiment Results**

Clearly, neither the “1 mile or less” and “Over 1 mile” groups, nor the “Constrained” and “Unconstrained” groups share similar consumption behavior. Endogenous sorting near grocery stores seems to have some influence on respondents’ consumption since compensation required (negative willingness-to-pay) for added travel time is significantly greater for those respondents who live nearby grocery stores. This result provides evidence in support of Hypothesis 4. However, both the “Over 1 mile” group and the “Constrained” group place relatively high value on high quality produce, which is inconsistent with their volatile consumption of fresh produce. Further, the “Constrained” group appears to value grocery and specialty stores highly, despite the fact that they patronize supercenters more frequently, another inconsistency. These results provide support for Hypothesis 5 – the exogenous food geography likely explains the behavioral differences between the groups because the “Constrained” group’s preferences appear at odds with behavior. Consequently, exogenous treatment effects also explain divergent consumption behavior, in addition to locational sorting.

Whether the exogenous treatment effects of grocery store proximity or locational sorting near grocery stores are more to blame for disparities in fresh produce consumption depends on further investigation of the “Constrained” and “Unconstrained” groups. I cannot conclude that either the “1 mile or less” and “Over 1 mile” groups share similar preferences, nor can I conclude that the “Constrained” and “Unconstrained” groups share similar preferences. However, comparing the differences in marginal WTP for these two sets suggests that the “Constrained” and “Unconstrained” groups have comparably dissimilar marginal WTP values for grocery store attributes as do the “1 mile or less” and “Over 1 mile” groups. Consumption behavior, on the other hand, contrasts more for the “Constrained” and “Unconstrained” groups than for the “1 mile or less” and “Over 1 mile” groups. The “Constrained” group, for example, exhibited highly elastic demand for fresh vegetables (see Table 12), and a high propensity of substituting fresh for canned or frozen produce (see Table 13). Thus, the “Constrained” group signaled in the choice experiment that they would prefer a grocery shopping scenario like that of the “Unconstrained” group, however their behavior exposes significant contrasts to their “Unconstrained”

counterparts. This result provides evidence in support of Hypothesis 6: exogenous treatment effects appear to explain consumption behavior among the survey respondents more so than locational sorting. In the following section, I conduct a supplementary analysis to investigate which specific aspects of the exogenous food environment can alter fresh produce consumption.

#### **5.3.4 Comparison of Constrained versus Willing Travelers**

I investigate whether quality, travel time, or some other factor might drive volatile fresh produce consumption by comparing the preferences of respondents who willingly travel further than they have to, the “Opt to Travel” respondents, against the “Constrained” group. For example, I can infer that lack of access to quality produce “constrains” the “Constrained” group if their marginal WTP for travel time is similarly low to that of the “Opt to Travel” group, but their marginal WTP for produce quality is significantly higher than that of the “Opt to Travel” group. Recall from Table 3 that the “Opt to Travel” respondents chose their status quo option 3.0 times out of six, on average, and the “Constrained” group chose their status quo option 1.1 times out of six on average. Both groups are relatively dissatisfied with their current grocery shopping situation compared to the full sample of respondents who chose their status quo option 3.3 times out of six. I investigate how their preferences and behavior compare.

Table 19 presents MMNL regression results for a comparison of the “Constrained” and “Opt to Travel” respondent groups. The dummy variable here is equal to 1 if the respondent falls in the “Opt to Travel” group and equal to 0 if the respondent falls in the “Constrained” group. Pooled regression results for these two groups are uninformative, thus I present only results of the model with the “Opt to Travel” dummy variable interacted with each main effect, following the previous MMNL result discussions. The “Opt to Travel” dummy is denoted by *OCD*.<sup>17</sup>

---

<sup>17</sup> Triple interaction terms with the *OCD* dummy were each insignificant, as with the previous models.

**Table 19: MMNL Results with "Opt to Travel" dummies**

	MMNL - Opt to Travel dummy		
Variable	Coef (SE)		SD Pval <sup>a</sup>
Main Effects			
ASC	-0.513	*	
	0.266		
Weekly Grocery Bill	-0.033		
	0.023		
Travel Time	-0.088	***	0.003
	0.022		
Specialty Store	0.630		0.000
	0.414		
Grocery Store	0.124		0.000
	0.300		
HQ Produce	3.448	***	0.000
	0.810		
AQ Produce	-0.198		0.279
	0.614		
Main Effect 'Opt to Travel' Dummies			
ASC_OCD	-2.263	***	0.000
	0.338		
Weekly Grocery Bill_OCD	-0.008		0.176
	0.020		
Travel Time_OCD	-0.047	*	0.001
	0.026		
Specialty Store_OCD	-0.685		0.069
	0.474		
Grocery Store_OCD	-0.394		0.004
	0.342		
HQ Produce_OCD	-0.650		0.907
	0.530		
AQ Produce_OCD	-0.259		0.357
	0.401		
Interaction Effects			
Distance x Grocery Bill	-0.009	*	0.796
	0.005		
CurrentTravel x Grocery Bill	-0.001		0.000
	0.001		
Income x HQ Produce	0.232	**	0.582
	0.110		
Income x AQ Produce	0.200	**	0.899
	0.091		
Observations			
Log Likelihood	5,346		
Prob > Chi2	-1244.49		
	0.000		
AIC	2524.98		
BIC	2643.50		

\*\*\* significant at 1%; \*\* significant at 5%, \* significant at 10%

<sup>a</sup>Significance of standard deviations of individual heterogeneity

With the exception of the travel time attribute, the “Constrained” and “Opt to Travel” groups share statistically similar utility for grocery store preferences. Both groups have disutility from a higher grocery bill based on the linear combination of estimators test for significance provided in Appendix F. Results of Table 20 illustrate each group’s marginal WTP for store attributes.

**Table 20: Marginal Willingness to Pay – comparison of constrained versus willing traveler respondents**

	Constrained	Opt to Travel	Significant Difference?*
Travel Time (\$/min. travel)	-\$1.10	-\$2.38	Yes
Specialty Store (\$/specialty vs supercenter)	\$7.86	-\$0.96	No
Grocery Store (\$/grocery vs supercenter)	\$1.55	-\$4.77	No
HQ Produce (\$/HQ vs LQ produce)	\$60.08	\$70.72	No
AQ Produce (\$/AQ vs LQ produce)	\$12.23	\$10.29	No
<b>Unit Values</b>			
Average Travel Time (min.)	12.3	12.1	
Average Distance (mi.)	4.0	0.4	
Average Income Bracket <sup>^</sup>	5.9	5.2	

\*Significant difference between sample groups of utility for the attribute, not marginal WTP for the attribute

<sup>^</sup>Income bracket of 4 denotes annual income between \$50-75K; 5 denotes \$75-100K; 6 denotes \$100-125K

As I expect, both groups require similar compensation for added travel time. This result confirms that the “Constrained” group respondents are not likely “constrained” by their longer trips to the grocery store because they are no more averse to added travel than people who willingly travel but do not need to.

The “Constrained” group has similar marginal WTP for high quality produce as does the “Opt to Travel” group. The difference of \$60.08 versus \$70.72 is small compared to previous respondent group comparisons. The “Constrained” group displayed robust tendency of choosing shopping alternatives with high quality produce (see Appendix D), which is inconsistent with the fact that they consumed five percent less fresh produce and almost five percent more non-fruit & vegetable “other foods” than did the “Opt to Travel Group” (see Table 7) and is also inconsistent with their more elastic demand for fresh produce (see Table 12 and Table 13). Thus, the “Constrained” group may be deterred from consuming fresh produce because they are unsatisfied with the quality available at their main grocery retailers. The difference in marginal WTP for high quality produce between the two groups most likely stems from familiarity differences with high quality produce.

Dissatisfaction with store type also appears to affect the “Constrained” group’s consumption patterns. The “Constrained” and “Opt to Travel” groups travel similar distances to get to their main grocery store. Note the average stated travel times are 12.3 and 12.1, respectively. Also, both groups have a relatively high incidence of shopping at supercenters in the month prior to the survey, as shown by Table 4. Yet the marginal WTP for store types suggest the “Opt to Travel” group actually prefers to shop at supercenters whereas the “Constrained” group would greatly prefer to shop at a specialty store. Consequently, produce quality and store amenities appear to have the strongest constraining effects when considering how the exogenous food environment can affect produce consumption.

#### 5.4 Results of Robustness Tests

In the following section, I run several robustness checks to test the sensitivity of my conclusions to my respondent group classifications. First I alter the geographic constraint classification from the USDA's 1-mile cut-off to a 2-mile cut-off. Second, I use only a lower-income sub-sample to test if the above-average income levels of my respondents bias my conclusions. Third, I expand the sample of "Constrained" respondents to test for any small-sample bias in the preference-based classification analyses. In each of these three robustness checks I mimic the methodology presented in sections 5.2 and 5.3 where I use a demand system and a choice experiment to compare behavior and preferences across respondent groups. My results are robust across each classification. I find proximity to a grocery store is positively related to stable fresh fruit & vegetable consumption and lower propensity to substitute non-fresh for fresh vegetables. I also find the choice experiment results to be robust, suggesting both exogenous and endogenous forces affect fresh produce consumption, however quality and store amenities possess a stronger constraining effect than do cost or travel time.

The 1-mile cut-off I use to differentiate respondents may be too restrictive, thus I alter the geographic constraint classification and consider respondents who live within two miles of a grocery store and compare their behavior and preferences to respondents who live over two miles from a grocery store. Of the respondents who provided neighborhood location data, 56 live over two miles from a grocery store, while 515 live within two miles of a grocery store. Table 21 compares the expenditure and own-price elasticities of "2 mile or less" respondents and "Over 2 mile" respondents. The "2 mile or less" respondents exhibit a relatively inelastic expenditure elasticity of demand for fresh vegetables of 0.85 compared to the "Over 2 miles" group who have an expenditure elasticity of 1.06. Expenditure elasticity of demand for fresh fruit is identical between the two groups.

Table 22 provides the cross-price elasticities of demand. The "Over 2 miles" respondents are more likely to substitute fresh for canned vegetables with a cross price elasticity of 0.15 compared to the cross price elasticity of 0.09 for the "2 miles or less" group.

**Table 21:** Estimated expenditure and own-price elasticities using a 2-mile cut-off

$e_I, e_p$	2 miles or less		Over 2 miles	
	Food Expenditure	Marshallian own-price	Food Expenditure	Marshallian own-price
<b>Food Group</b>				
(1) Fresh Veg	0.85	-0.82	1.06	-1.08
(2) Fresh Fruit	0.76	-0.83	0.76	-0.77
(3) Canned Veg	1.05	-0.38	1.29	-0.96
(4) Canned Fruit	0.99	-0.41	1.19	-0.21
(5) Frozen Veg	0.89	-1.19	1.14	-1.06
(6) Frozen Fruit	0.85	1.12	1.20	-0.43
(7) Other Foods	1.03	-0.99	1.00	-1.01
Observations	515		56	

**Table 22:** Estimated Marshallian cross-price elasticities using a 2-mile cut-off

$e_{x, p(y)}$	2 miles or less							Over 2 miles						
	Fresh Veg	Fresh Fruit	Canned Veg	Canned Fruit	Frozen Veg	Frozen Fruit	Other Foods	Fresh Veg	Fresh Fruit	Canned Veg	Canned Fruit	Frozen Veg	Frozen Fruit	Other Foods
<b>Food Group</b>														
(1) Fresh Veg	<b>-0.82</b>	0.06	0.17	0.14	0.12	0.16	0.24	<b>-1.08</b>	0.00	-0.05	0.00	-0.02	0.02	0.10
(2) Fresh Fruit	-0.02	<b>-0.83</b>	0.20	0.21	0.14	0.20	0.40	0.04	<b>-0.77</b>	-0.03	-0.10	-0.15	-0.04	-0.36
(3) Canned Veg	0.09	0.05	<b>-0.38</b>	-0.40	0.23	-0.50	-0.08	0.15	0.03	<b>-0.96</b>	0.20	-0.16	-0.01	0.44
(4) Canned Fruit	-0.14	0.19	-1.44	<b>-0.41</b>	0.34	0.42	0.03	-0.03	-0.11	-0.19	<b>-0.21</b>	-0.33	-0.55	0.23
(5) Frozen Veg	-0.26	-0.17	0.46	0.21	<b>-1.19</b>	-0.06	0.10	-0.06	-0.14	0.28	-0.27	<b>-1.06</b>	-0.02	0.20
(6) Frozen Fruit	0.30	0.26	-3.21	0.77	-0.24	<b>1.12</b>	0.04	0.11	0.00	0.34	-1.27	-0.04	<b>-0.43</b>	0.21
(7) Other Foods	0.04	0.04	0.03	0.03	0.03	0.03	<b>-0.99</b>	0.02	-0.02	-0.02	0.01	0.01	0.00	<b>-1.01</b>
Observations	515							56						

Considering the preferences of these two groups, Table 23 provides marginal WTP calculations. These marginal WTP results mimic previous results under the 1-mile geographic classification shown in Table 16. Respondents who live over two miles from a grocery store require less compensation for travel time and are less willing-to-pay for both high and average quality produce compared to the “2 miles or less” group. One significant change from the 2-mile cut-off is that proximate respondents are more willing to pay to shop at a specialty or grocery store instead of a supercenter, whereas the “Over 2 miles” group would prefer to shop at a supercenter. MMNL regression results using a 2-mile cut-off are provided in Appendix G.

**Table 23: Marginal Willingness to Pay - 2 mile cut-off**

	2 miles or less	Over 2 miles	Significant Difference?*
Travel Time (\$/min. travel)	-\$3.06	-\$1.59	No
Specialty Store (\$/specialty vs supercenter)	\$12.28	-\$4.90	Yes
Grocery Store (\$/grocery vs supercenter)	\$0.91	-\$3.20	No
HQ Produce (\$/HQ vs LQ produce)	\$82.84	\$57.06	Yes
AQ Produce (\$/AQ vs LQ produce)	\$19.75	\$16.33	No
<b>Unit Values</b>			
Average Travel Time (min.)	8.8	10.2	
Average Distance (mi.)	0.6	7.1	
Average Income Bracket <sup>^</sup>	5.4	5.4	

\*Significant difference between sample groups of utility for the attribute, not marginal WTP for the attribute

<sup>^</sup>Income bracket of 4 denotes annual income between \$50-75K; 5 denotes \$75-100K; 6 denotes \$100-125K

The 2-mile cut-off results appear robust with the 1-mile cut-off results. Proximity is related to relatively inelastic demand for fresh vegetables, and less propensity to substitute non-fresh for fresh vegetables. The choice experiment results suggest locational sorting may drive some of these behavioral differences, yet lack of access to quality produce may also drive some of these behavioral differences.

Next, I consider only survey respondents with annual household income less than \$75,000 per year to test whether the above-average income-levels of my respondents have biased my results. The average annual household income of this lower-income subsample is approximately \$42,500, which is more representative of the national “food desert” resident population as a whole.<sup>18</sup> Of the “1 mile or less” group, 34.3 percent (156 respondents) have lower income and of the “Over 1 mile” group, 18.1 percent (21 respondents) have lower income.

Table 24 and Table 25 provide expenditure, own-price and cross-price demand elasticities comparing “1 mile or less” respondents with “Over 1 mile” respondents. I aggregated the food expenditure categories into five broad groups in order to estimate a demand system for the smaller sample of respondents. “FC Veg” stands for frozen and canned vegetables and “FC Fruit” stands for frozen and canned fruit. Results of the lower-income sub-sample demand system mimic previous results. Table 24 shows the respondents who live closer to a grocery store have relatively inelastic demand for all fruits and vegetables compared to those who live over one mile from a grocery store; however the lower-income “Over 1 Mile” group treats “Other Foods” as a staple item more so than the proximate respondents. Thus proximity to a grocery store is correlated with more stable produce consumption and less reliance on non-fruit and vegetable foods for the full sample, as well as for lower-income respondents.

<sup>18</sup> Median household income among “food desert” residents was \$35,860.64 based on the American Community Survey 2005-09 (Dutko, Ver Ploeg, and Farrigan 2012).

Table 25 also reveals respondents who live over one mile from a grocery are significantly more likely to substitute non-fresh for fresh vegetables with a cross-price elasticity of 0.28 compared to those who live within one mile of a grocery who appear to treat fresh and non-fresh vegetables as complementary goods with a cross-price elasticity of -0.18.

**Table 24:** Estimated expenditure and own-price elasticities of lower-income respondent sub-sample

$e_I, e_p$	Lower Income - 1 Mile or Less		Lower Income - Over 1 Mile	
	Food	Marshallian	Food	Marshallian
	Expenditure	own-price	Expenditure	own-price
Food Group				
(1) Fresh Veg	0.91	-0.92	1.33	-0.98
(2) Fresh Fruit	0.84	-0.92	1.09	-1.10
(3) FC Veg	0.93	-0.47	1.41	-1.54
(4) FC Fruit	0.93	-0.59	1.22	-1.35
(5) Other Foods	1.02	-0.98	0.60	-1.02
Observations	156		21	

**Table 25:** Estimated Marshallian cross-price elasticities of lower-income respondent sub-sample

$e_{x, p(y)}$	Lower Income - 1 Mile or less					Lower Income - Over 1 Mile				
	Fresh	Fresh			Other	Fresh	Fresh			Other
	Veg	Fruit	FC Veg	FC Fruit	Foods	Veg	Fruit	FC Veg	FC Fruit	Foods
Food Group										
(1) Fresh Veg	<b>-0.92</b>	0.14	0.06	0.11	0.09	<b>-0.98</b>	0.19	0.04	-0.16	0.11
(2) Fresh Fruit	0.16	<b>-0.92</b>	0.06	0.11	0.20	-0.13	<b>-1.10</b>	-0.09	0.12	-0.01
(3) FC Veg	-0.18	-0.12	<b>-0.47</b>	-0.19	<b>0.05</b>	0.28	0.34	<b>-1.54</b>	0.11	0.05
(4) FC Fruit	0.22	-0.04	-0.52	<b>-0.59</b>	-0.04	-0.07	-0.07	-0.18	<b>-1.35</b>	-0.14
(5) Other Foods	0.03	0.03	0.03	0.03	<b>-0.98</b>	-0.23	-0.02	0.03	0.32	<b>-1.02</b>
Observations	156					21				

Choice experiment results for the lower-income sub-sample are robust with the full sample of respondents. Table 26 shows that those who live over one mile of a grocery store have a higher marginal WTP to shop at either a grocery or specialty store instead of a supercenter despite the fact that this sub-sample of 21 respondents was more likely to have shopped at a supercenter in the month prior to the survey compared to the “1 mile or less” lower-income group.<sup>19</sup> Also, both groups have similarly high marginal WTP for produce quality. These results, again, suggest the behavioral disparity particularly of fresh vegetable consumption may be driven in part by exogenous constraints to high quality produce or smaller, specialized grocery stores. MMNL regression results based on the lower-income sub-sample are provided in Appendix G.

<sup>19</sup> Of the 21 respondents comprising the lower-income “Over 1 mile” group, 76.2 percent (16 respondents) had shopped at a supercenter in the month prior to the survey. Of the 156 respondents comprising the lower-income “1 mile or less” group, 58.3 percent (91 respondents) had shopped at a supercenter in the month prior to the survey.



**Table 26:** Marginal Willingness to Pay - Lower income respondent sub-sample

	1 mile or less	Over 1 mile	Significant Difference?*
Travel Time (\$/min. travel)	-\$1.87	-\$1.11	No
Specialty Store (\$/specialty vs supercenter)	\$5.61	\$8.43	No
Grocery Store (\$/grocery vs supercenter)	-\$4.25	\$3.13	No
HQ Produce (\$/HQ vs LQ produce)	\$48.51	\$35.38	No
AQ Produce (\$/AQ vs LQ produce)	\$12.21	\$5.29	No
<b>Unit Values</b>			
Average Travel Time (min.)	9.2	9.9	
Average Distance (mi.)	0.4	4.1	
Average Income Bracket^	3.5	3.5	

\*Significant difference between sample groups of utility for the attribute, not marginal WTP for the attribute

^Income bracket of 3 denotes annual income between \$35-50K; 4 denotes \$50-75K

The third robustness test increases the sample of “Constrained” respondents to test for any small-sample bias in the preference-based classification analyses. Recall, the “Constrained” respondents are a subset of the “Over 1 mile” group who have further signaled dissatisfaction with their current grocery shopping situation.

“Unconstrained” respondents, alternatively, are a subset of the “1 mile or less” group who signaled above-average satisfaction with their current grocery shopping situation.

Originally I categorized “Constrained” as respondents who live over 1 mile from a grocery store and selected their status quo grocery store as opposed to a hypothetical grocery store option less than 3 times out of 6, or less than the average. This resulted in a very small sample of 56 respondents. I expand this group to include respondents who chose their status quo 3 times or less out of 6 and check for significant differences when comparing their preferences and behavior to that of the “Unconstrained” respondents. With the broader definition, the number of “Constrained” respondents changed from 56 to 82. This updated definition for the “Constrained” group does not substantially alter the results. The demand estimation results in Table 27 and Table 28, again, support my previous conclusions that proximity to a grocery store is related to greater consistency of fresh produce consumption. The “Constrained” group is more likely to substitute canned for fresh vegetables and canned for fresh fruit compared to the “Unconstrained” group.

**Table 27:** Estimated expenditure and own-price elasticities using a larger "Constrained" respondent group

Food Group	$e_I, e_p$		Unconstrained		Constrained	
	Expenditure	Marshallian own-price	Food Expenditure	Marshallian own-price	Food Expenditure	Marshallian own-price
(1) Fresh Veg			0.33	-0.35	0.90	-0.82
(2) Fresh Fruit			0.70	-0.88	0.99	-0.90
(3) Canned Veg			1.15	-0.90	1.01	0.07
(4) Canned Fruit			1.16	-1.04	1.21	-1.70
(5) Frozen Veg			1.15	-1.01	1.11	-0.91
(6) Frozen Fruit			1.20	-1.01	1.20	-1.02
(7) Other Foods			1.02	-0.99	1.00	-0.99
Observations			221		81	

**Table 28:** Estimated Marshallian cross-price elasticities using a larger "Constrained" respondent group

Food Group	$e_{x,p(y)}$		Unconstrained							Constrained						
	Fresh Veg	Fresh Fruit	Canned Veg	Canned Fruit	Frozen Veg	Frozen Fruit	Other Foods	Fresh Veg	Fresh Fruit	Canned Veg	Canned Fruit	Frozen Veg	Frozen Fruit	Other Foods	Fresh Veg	Fresh Fruit
(1) Fresh Veg	<b>-0.35</b>	0.31	0.36	0.33	0.28	0.49	-0.29	<b>-0.82</b>	0.05	0.06	-0.14	-0.01	0.11	-0.10		
(2) Fresh Fruit	-0.12	<b>-0.88</b>	0.05	0.01	0.05	0.11	-0.14	0.03	<b>-0.90</b>	-0.02	-0.09	-0.05	0.07	-0.03		
(3) Canned Veg	0.09	0.03	<b>-0.90</b>	0.13	-0.10	-0.32	0.08	0.10	-0.08	<b>0.07</b>	-3.78	1.25	1.45	-0.01		
(4) Canned Fruit	0.07	-0.05	0.37	<b>-1.04</b>	-0.02	-0.42	0.12	0.10	0.07	0.26	<b>-1.70</b>	0.08	0.30	0.16		
(5) Frozen Veg	0.00	0.03	-0.10	-0.01	<b>-1.01</b>	0.00	0.10	-0.05	-0.12	0.69	-0.44	<b>-0.91</b>	-0.27	0.09		
(6) Frozen Fruit	0.13	0.04	-0.15	-0.06	0.01	<b>-1.01</b>	0.12	0.00	-0.01	0.03	-0.29	-0.07	<b>-1.02</b>	0.09		
(7) Other Foods	-0.04	-0.02	-0.01	-0.01	-0.01	0.00	<b>-0.99</b>	-0.02	-0.01	0.00	-0.08	0.01	0.09	<b>-0.99</b>		
Observations		221													81	

The choice experiment results using the larger "Constrained" sample are robust to previous results shown in Table 18. Results of the MMNL regression incorporating the broad "Constrained" definition are provided in Appendix G. From Table 29, the larger "Constrained" group appears, again, to be dissatisfied with the amenities of their current grocery retailers. They would prefer to shop at a specialty store or a grocery store instead of a supercenter despite the fact that this group is more likely to have shopped at a supercenter in the month prior to the survey, compared to the "Unconstrained" group. The incidence of shopping at a supercenter continues to be high relative to the other respondent groups, even after broadening the "Constrained" respondent definition.<sup>20</sup> Thus, it appears the "Constrained" group would prefer to shop at a store type different from a supercenter, but geographic constraints prevent them from doing so.

Travel time is not likely to be the constraining factor altering the "Constrained" group's consumption behavior. The "Constrained" group has a lower marginal WTP for added travel than the "Unconstrained" group, and the coefficient on *Travel Time\_CD* is positive meaning the "Constrained" group has less disutility from added travel (see Appendix G). Also, income is not likely driving the larger "Constrained" group to shop at supercenters, since they have higher average income than the "Unconstrained" group. While The "Constrained" group has lower

<sup>20</sup> Of the 81 respondents in the larger "Constrained" group, 85.2 percent (69 respondents) had shopped at a supercenter in the month prior to the survey compared to 48.4 percent for the "Unconstrained" group (see Table 4).

marginal WTP for high quality produce, some of this result may be driven by each group's respective experience with "high quality" produce. Note the "Constrained" group is willing to pay more for average quality produce than the "Unconstrained." Further, the "Constrained" group garners statistically similar utility for high quality produce as does the "Unconstrained" group, since the coefficient for the high quality parameter estimates are not statistically distinct for these groups. The differences in purchasing behavior and status quo satisfaction between the "Constrained" versus "Unconstrained" groups may result from differences in each groups' average grocery shopping experience including ambience, layout, or product offerings, as opposed to added travel or price sensitivity.

**Table 29: Marginal Willingness to Pay - larger "Constrained" respondent group**

	Unconstrained	Constrained	Significant Difference?*
Travel Time (\$/min. travel)	-\$3.87	-\$0.86	Yes
Specialty Store (\$/specialty vs supercenter)	\$0.82	\$14.39	Yes
Grocery Store (\$/grocery vs supercenter)	-\$4.68	\$6.18	Yes
HQ Produce (\$/HQ vs LQ produce)	\$95.97	\$59.52	No
AQ Produce (\$/AQ vs LQ produce)	\$8.31	\$11.98	No
<b>Unit Values</b>			
Average Travel Time (min.)	7.4	13.1	
Average Distance (mi.)	0.5	4.2	
Average Income Bracket <sup>^</sup>	5.6	5.8	
<i>Observations</i>	<i>221</i>	<i>82</i>	

\*Significant difference between sample groups of utility for the attribute, not marginal WTP for the

<sup>^</sup>Income bracket of 4 denotes annual income between \$50-75K; 5 denotes \$75-100K; 6 denotes \$100-125K

## **Chapter Six: Summary of Findings and Conclusions**

The objective of this thesis was to identify a relationship between food geography and consumption behavior. My identification strategy involved examining the discrete choice of where to grocery shop against the continuous choice of which food products to purchase at the grocery store. I used a non-market valuation method known as a choice experiment to quantify value perception of grocery store attributes, and I estimated a system of demand for grocery items using original survey data from 991 residents of Champaign & Piatt counties, IL.

This research was motivated by growing awareness of the limitations of the “food desert” construct both in the media, and in the academic community. Current literature on food access and food geography has demonstrated that a greater number of supermarkets in a finite geographic area are positively correlated with greater health for residents of that area, but causality remains inconclusive. Literature has also found that “food desert” residents have relatively elastic demand for fresh produce, however it remains undetermined whether such demand behavior is a direct result of food geography, or a result of inherent preferences among “food desert” residents.

The methodology described in this thesis contributes to the current body of literature on food demand and food access in several ways. First, I address the simultaneous choice problem inherent in grocery shopping in which the choice of where to grocery shop has direct implications on what foods one decides to purchase. Second, I provide an approach for separating endogenous sorting from exogenous treatment effects in order to isolate how the food environment can influence fruit & vegetable consumption.

I find proximity to a grocery store is positively related to stable consumption of fresh produce and decreases the propensity to substitute fresh for non-fresh vegetables. Lack of access to specialized, small-scale retailers, as well as higher quality produce appears to be the most significant drivers of price elastic produce consumption and a higher frequency of trading off non-fresh for fresh produce. Neither added travel time nor expense appears to affect consumption behavior among survey respondents. Previous literature and the results of summary statistics from this thesis’ sample data imply satisfaction with one’s grocery store and a higher perception of quality value are positively related to fruit & vegetable consumption. Thus travel time and geographic distance in isolation appear less likely to alter produce consumption than does access to quality products and quality retailers.

I further find that proximity to a grocery store is positively related to consumption of fresher foods even for respondents who travel further than necessary to grocery shop. Respondents who willingly traveled over one mile to grocery shop exhibited similar preferences as respondents who were forced to travel over one mile to grocery shop. However, the respondents who opt to travel consumed fresh produce to a greater extent. This finding suggests that the food environment may have some indirect spillover effects that cause nearby residents to choose fresh foods more often than non-fresh foods. Alternatively, this finding suggests that an exogenous food desert

definition based on geographic boundaries overstates the influence of the food environment, thus overstates the potential constraining effects of the food environment.

Both endogenous sorting and environmental treatment effects of the food environment influence food consumption. While respondents who live within one mile of a grocery store consumed greater proportions of, and had inelastic demand for, fresh produce compared to those who live over one mile from a grocery store, these proximate respondents also exhibited stronger preferences for proximity and high quality produce. However, treatment effects appear to exercise greater influence on consumption behavior than does endogenous sorting among survey respondents.

Current policy response to inadequate food access focuses first on a geographic, exogenous definition to define areas as “food deserts,” and second encourages grocery chains to enterprise in underserved areas. The results of this thesis suggest policy makers should equally focus on quality improvements such as advancing store amenities and quality of produce offered in stores. Addressing quality before quantity may provide a more effective means of encouraging fresh produce consumption of nearby residents. Altering the food environment will not provide a comprehensive solution to encouraging fresh produce consumption in underserved areas. Alternative means of health intervention, such as nutrition education are likely to be necessary for populations of people for whom personal preferences are stronger determinants of consumption behavior than their exogenous food environment.

The methods suggested in this thesis are not without limitations. I do not have a statistically sound technique for concluding whether results of the demand system are consistent with preferential results of the choice experiment analyses. My conclusions are specific to a sample of people who are not representative of the US as a whole, thus I cannot generalize my findings beyond the sample of residents of Champaign & Piatt counties affiliated with the University of Illinois. The high fresh fruit & vegetable consumption across all sample groups is biased due to the higher income and education levels of my sample (Blisard, Stewart, and Joliffe 2004). Repeating the analyses presented in this paper with a lower-income population may result in different conclusions. Lastly, I am not able to purely isolate treatment effects from preference endogeneity in absence of a clear “before and after” experiment involving a new grocery in Champaign County. However, the methods introduced in this thesis have the potential to focus and improve policy and improve the health and well-being of underserved individuals faced with inescapable barriers to health.

## References

- Abdulai, Awudu (2002). "Household Demand for Food in Switzerland: A Quadratic Almost Ideal Demand System." *Swiss Society of Economics and Statistics*, 138(1): 1-18.
- Alpizar, Francisco, Fredrik Carlsson, and Peter Martinsson. 2001. "Using Choice Experiments for Non-Market Valuation." Working paper. Göteborg University Department of Economics.
- Andreyeva, Tatiana, Michael Long, and Kelly Brownell. 2010. "The Impact of Food Prices on Consumption: A Systematic Review of Research on the Price Elasticity of Demand for Food." *American Journal of Public Health*, 100(2): 216-222.
- Arnold, Stephen J, Tae H Oum & Douglas J. Tigert. 1983. "Determinant Attributes in Retail Patronage: Seasonal, Temporal, Regional, and International Comparisons." *Journal of Marketing Research* Vol. 20 (May 1983): 149-57.
- Banks, James, Richard Blundell, and Arthur Lewbel. (1997). "Quadratic Engel Curves and Consumer Demand." *The Review of Economics and Statistics*, 79(4): 527-539.
- Bayer, Patrick and Christopher Timmins. 2005. "On the equilibrium properties of locational sorting models." *Journal of Urban Economics* 57(2005): 462-477.
- Becker, Gary. 1965. "A Theory of the Allocation of Time." *Economic Journal*. 75:493-517.
- Bentham, Jeremy. 1789. *Introduction to the Principles of Morals and Legislation*. Oxford: Clarendon Press.
- Berg, Nathan and James Murdoch. 2007. "Access to Grocery Stores and Food Security in Dallas." Research Brief. The J. McDonald Williams Institute Center for Urban Economics. University of Texas – Dallas.
- Bernstein, David. 2012, April 25. "Time to Revisit Food Deserts." *The New York Times*. Retrieved from <http://opinionator.blogs.nytimes.com/2012/04/25/time-to-revisit-food-deserts/>
- Bitler, Marianne and Steven J. Halder. 2011. "An Economic View of Food Deserts in the United States." *Journal of Policy Analysis and Management*. 30(1): 153-176.
- Blisard, Noel, Hayden Stewart, and Dean Jolliffe. 2004. *Low-Income Households' Expenditures on Fruits and Vegetables*. Washington DC: US Department of Agriculture, ERS Agricultural Economic Report Number 833, May.
- Blitstein, Jonathan, Jeremy Snider, and W Douglas Evans. 2011. "Perceptions of the food shopping environment are associated with greater consumption of fruits and vegetables." *Public Health Nutrition* 15(6): 1124-1129.
- Brown, Daniel. 1978. "An Examination of Consumer Grocery Store Choice: Considering the Attraction of Size and the Friction of Travel Time." *NA – Advances in Consumer Research*. Vol 5.
- Brown, Stephen. 1989. "Retail Location Theory: The Legacy of Harold Hotelling," *Journal of Retailing* 65(4): 450-470.
- California Center for Public Health Advocacy. 2007. *Searching for Healthy Food: The Food Landscape in California Cities and Counties*. Davis, CA.

- Carlson, Andrea, Mark Lino, WenYen Juan, Kenneth Hanson, and P. Peter Basiotis. 2007. *Thrifty Food Plan, 2006*. Washington DC: US Department of Agriculture, ERS CNPP-19, April.
- Carrasco, N. (2008) Deciding where to shop: disaggregate random utility destination choice modeling of grocery shopping in canton Zurich, *Master Thesis*, IVT, ETH Zurich, Zurich.  
<[www.ivt.ethz.ch/docs/students/sa223.pdf](http://www.ivt.ethz.ch/docs/students/sa223.pdf)>
- Centers for Disease Control & Prevention. National Center for Chronic Disease Prevention and Health Promotion. Division of Nutrition, Physical Activity, and Obesity. "State Initiatives Supporting Healthier Food Retail: An Overview of the National Landscape." <[www.cdc.gov/obesity/downloads/Healthier\\_Food\\_Retail.pdf](http://www.cdc.gov/obesity/downloads/Healthier_Food_Retail.pdf)>
- Chen, S., Florax, R.J.G.M., and Snyder, S.D. (2010). Does where you live make you fat? Obesity and access to chain grocers. *Economic Geography*, 86, 431-452.
- Currie, J., DellaVigna, S., Moretti, E., and Pathania, V. (2009). The effect of fast food restaurants on obesity and weight gain. *NBER Working Paper No. 14721*, Retrieved March 2012 from  
<<http://www.nber.org/papers/w14721>>.
- Deaton, Angus (1988). "Quality, quantity, and spatial variation of price." *American Economic Review*, 78(3): 418-430.
- Dissanayake, Sahan and Amy Ando (2011). "Valuing restoration: conservation tradeoffs and respondent experience." Working paper.
- Deaton, Angus and John Muellbauer. 1980a. "An Almost Ideal Demand System." *American Economic Review*, Vol. 70, 899-936.
- Deaton, Angus and John Muellbauer. 1980b. *Economics and Consumer Behavior*. Cambridge University Press.
- "Detroit Grocery Stores, 'Food Desert' Myth Examined in Noah Stephens' Photography Survey." Huffington Post. 22 October 2012. [http://www.huffingtonpost.com/2012/10/22/detroit-grocery-stores-food-desert-noah-stephens\\_n\\_2001896.html](http://www.huffingtonpost.com/2012/10/22/detroit-grocery-stores-food-desert-noah-stephens_n_2001896.html). Viewed 16 February 2013.
- Drewnoski, A (2003). "Fat and sugar: An economic analysis" *Journal of Nutrition*, 133, 838S-840S.
- Drewnowski, A., Darmon, N., & Briend, A. (2004). "Replacing fats and sweets with vegetables and fruit – a question of cost." *American Journal of Public Health*, 94, 1555-1559.
- Durham, Catherine and James Eales. (2010) "Demand elasticities for fresh fruit at the retail level." *Applied Economics*, 42: 1345-1354.
- Dutko, Paul, Michele Ver Ploeg, and Tracey Farrigan. 2012. *Characteristics and Influential Factors of Food Deserts*. Washington DC: US Department of Agriculture, ERS ERR-140, August.
- Eales, James S. and Laurian J Unnevehr. 1994. "The inverse almost ideal demand system." *European Economic Review*, 38:101-115.
- Erdil, Erkan. 2005. "Demand systems for agricultural products in OECD countries." *Applied Economics Letters*, 13:163-169.
- Gibson, John and Scott Rozelle. (2005). "Prices and Unit Values in Poverty Measurement and Tax Reform Analysis." *World Bank Economic Review* 19(1): 69-98.

- Gibson, D.M. (2011). The neighborhood food environment and adult weight status: estimates from longitudinal data. *American Journal of Public Health*, 101: 71.
- Giskes, K., van Lenthe, F., Avendano-Pabon, M., and Brug, J. (2011). A systematic review of environmental factors and obesogenic dietary intakes among adults: Are we getting closer to understanding obesogenic environments? *Obesity Review*, 12: e95-e106.
- Gómez, Miguel, Edward McLaughlin, and Dick Wittink. 2004. "Customer satisfaction and retail sales performance: an empirical investigation." *Journal of Retailing*. 80(2004): 265 – 278.
- Gorman, W.M. (1981). "Some Engel Curves" , in A. Deaton (ed). *Essays in the Theory and Measurement of Consumer Behaviour in Honour of Sir Richard Stone*, Cambridge.
- Greene, William. (a) Lecture notes. "Discrete Choice Modeling." Discrete Choice Modeling. New York University. <[https://files.nyu.edu/mrg217/public/unordered\\_multiresponse.pdf](https://files.nyu.edu/mrg217/public/unordered_multiresponse.pdf)> Visited 24 February 2013.
- Greene, William. (b) Lecture notes. "Discrete Choice Models I." Discrete Choice Modeling. New York University. <<http://pages.stern.nyu.edu/~wgreene/DiscreteChoice.htm>> Visited 3 March 2013.
- Hanemann, W.M. (2006). The economic conception of water, In: *Water Crisis: myth or reality?* Eds. P.P. Rogers, M.R. Llamas, L. Martinez-Cortina, Taylor & Francis plc., London.
- Hensher, Rose, Greene. *Applied Choice Analysis: A Primer* (2005). Cambridge University Press.
- Heiss, Florian. 2002. "Structural choice analysis with nested logit models." *The Stata Journal*. 2(3): 227-252.
- Herrmann, Roland and Claudia Roeder. 1998. "Some neglected issues in food demand analysis: retail-level demand, health information and product quality." *The Australian Journal of Agricultural and Resource Economics*. 42(4): 341 – 367.
- Hoch, Stephen, Byung-Do Kim, Alan Montgomery, and Peter Rossi. 1995. "Determinants of Store-Level Price Elasticity." *Journal of Marketing Research* 32(1): 17 –29.
- Hsieh, Ming-Feng and Kyle W. Stiegert. 2011. "Store Format Choice in Organic Food Consumption." *American Journal of Agriculture Economics*. 94(2):307-313.
- Huang, Kuo S. and Giing-Hwan Lin (200). "Estimation of Food Demand and Nutrient Elasticities from Household Survey Data." Economic Research Service, U.S. Department of Agriculture. Technical Bulletin No. 1887.
- Huang, Yan and Harmen Oppewal (2006), "Why consumers hesitate to shop online: An experimental choice analysis of grocery shopping and the role of delivery fees", *International Journal of Retail & Distribution Management*, Vol. 34.4:334-53.
- Izumi, Betty T., Shannon N. Zenk, Amy J. Schulz, Graciela B. Mentz, Sharon L. Sand, Ricardo F. de Majo, Christine Wilson, and Angela Odoms-Young. 2012. "Inter-Rater Reliability of the Food Environment Audit for Diverse Neighborhoods (FEAD-N)." *Journal of Urban Health* 89(3): 486 – 499.
- Katchova, Ani L. and Wen S. Chern (2004). "Comparison of Quadratic Expenditure System and Almost Ideal Demand System Based on Empirical Data." *International Journal of Applied Economics*, 1(1): 55-64.
- Kuhfeld, Warren. 2010. "Marketing Research Methods in SAS: Experimental Design, Choice, Conjoint, and Graphical Techniques," SAS 9.2 Edition MR-2010, 1 October 2010.

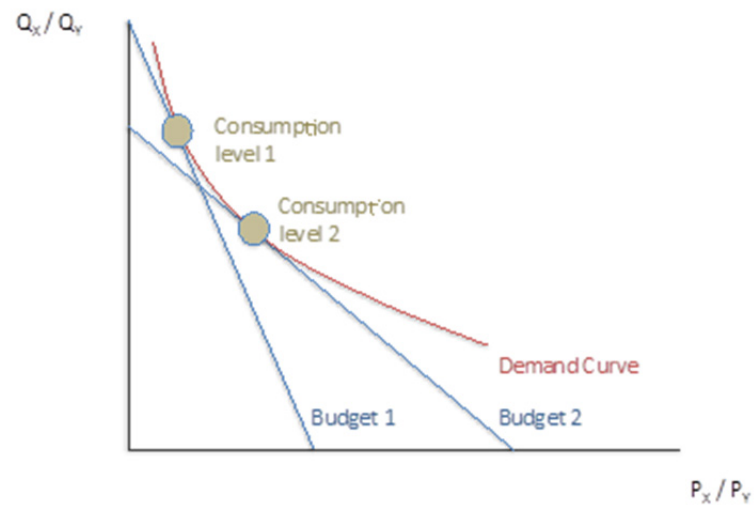


- Lancaster, Kelvin. (1966) "A New Approach to Consumer Theory." *The Journal of Political Economy*. 74(2): 132-157.
- Larson N, Story M, and Nelson M. (2009), "Neighborhood environments: disparities in access to healthy foods in the U.S.", *American Journal of Preventative Medicine* 2009, 36: 74-81.
- Lee, Jonq-Ying, Mark G. Brown and James L. Seale, Jr. 1994. "Model Choice in Consumer Analysis: Taiwan, 1970-89." *American Journal of Agricultural Economics*, 76:504-512
- Lewbel, Arthur. (1990). "Full Rank Demand Systems." *International Economic Review*, 31(2): 289-300.
- Liese, Angela D. Kristina E. Weis, Delores Pluto, Emily Smith, et al. 2007. "Food Store Types, Availability, and Cost of Foods in a Rural Environment." *Journal of the American Dietetic Association* 107:1916-1923.
- Louviere, J.J., D.A. Hensher and J.D. Swait, (2000). *Stated Choice Methods: Analysis and Applications*. Cambridge University Press.
- Mari Gallagher Research & Consulting Group. 2006. *Examining the Impact of Food Deserts on Public Health in Chicago*. Chicago, IL.
- McKelvey, Christopher. 2011. "Price, unit value, and quantity demanded." *Journal of Development Economics*. 95(2011): 157-169.
- Monsivais, Pablo and Adam Drewnowski (2007). "The Rising Cost of Low-Energy-Density Foods," *Journal of the American Dietetic Association* 107(12): 2071-2076.
- Morland K, Wing S, Diez-Roux A, Poole C. "Neighborhood characteristics associated with the location of food stores and food service places." *American Journal of Preventative Medicine*. 2002 (22): 23-29.
- Moro, Daniele and Paolo Sckokai (2000). "Heterogeneous preferences in household food consumption in Italy." *European Review of Agricultural Economics*, 27(3): 305-323.
- Myers, James and Mark Alpert. 1968. "Determinant Buying Attitudes: Meaning and Measurement." *Journal of Marketing*, 32 (October): 13-20.
- Nelson, Charles. Class Lecture Notes. Demand, Supply, Firms and Households. University of Illinois, Urbana-Champaign, IL. Spring 2013.
- Newby, P. K. 2006. "Examining Energy Density: Comments on Diet Quality, Dietary Advice, and the Cost of Healthful Eating." *Journal of the American Dietetic Association*, 106(8): 1166-1169.
- Nicholson, Walter and Christopher Snyder. 2008. *Microeconomic Theory: Basic Principles and Extensions*, 10<sup>th</sup> ed. Mason, OH: Thomson South-Western.
- Pearson, Tim, Jean Russell, Michael J Campbell, and Margo E Barker. 2005. "Do 'food deserts' influence fruit and vegetable consumption? – a cross-sectional study." *Appetite*, 45: 195-197.
- Raper, Kellie Curry, Maria Namakhoye Wanzala, and Rodolfo M. Nayga, Jr. 2002. "Food expenditures and household demographic composition in the US: a demand systems approach." *Applied Economics*, 34: 981-992.
- Rose, Donald and Rickelle Richards. 2003. "Food store access and household fruit and vegetable use among participants in the US Food Stamp Program." *Public Health Nutrition*, 7: 1081-1088.

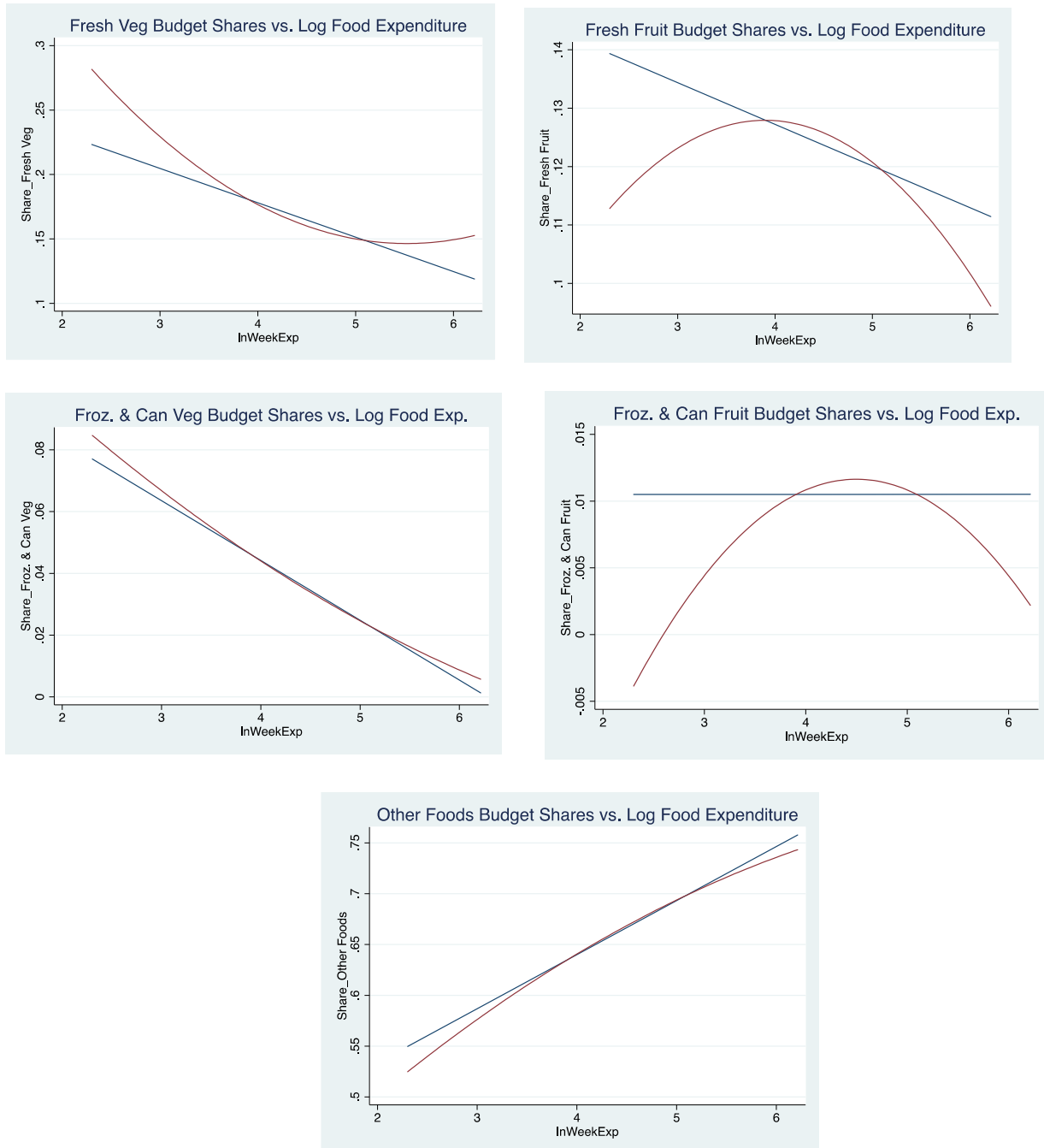
- Tang, Christopher S, David R Bell, Teck-Hua Ho (2001), "Store choice and shopping behavior: How price format works," *California Management Review* 43(2): 56-74.
- Teklu, Tesfaye and S.R. Johnson. 1986. "A Review of Consumer Demand Theory and Food Demand Studies in Indonesia." Staff Report 86-SR7. Center for Agricultural and Rural Development, Iowa State University.
- Tiebout, Charles. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64: 416-424.
- Tiffin, Richard, Kelvin Balcombe, Matthew Salois, and Ariane Kehlbacher. 2011. "Estimating Food and Drink Elasticities." University of Reading.
- U.S. Department of Health & Human Services. *Dietary Guidelines for Americans*. 2005. Washington DC <<http://www.health.gov/dietaryguidelines/dga2005/document/html/chapter5.htm>>
- Ver Ploeg, Michele, Vince Breneman, Tracey Farrigan, Karen Hamrick, David Hopkins, Phillip Kaufman, Biing-Hwan Lin, Mark Nord, Travis A. Smith, Ryan Williams, Kelly Kinnison, Carol Olander, Anita Singh, and Elizabeth Tukermanty. 2009. *Access to Affordable and Nutritious Food - Measuring and Understanding Food Deserts and Their Consequences: Report to Congress*. Washington DC: US Department of Agriculture, ERS AP-036, June.
- Weatherspoon, Dave, James Oehmke, Assa Dembele, Marcus Colema, Thasanee Satmanon, and Lorraine Weatherspoon. 2013. "Price and Expenditure Elasticities for Fresh Fruits in an Urban Food Desert." *Urban Studies*. 50(1): 88-106.
- Wilde, Park, Craig Gundersen, Katherine Baylis & Joseph Llobrera. "National Food Study Geography Component", Updated Study Plan, November 2012.
- Wilde, Park, Paul McNamara, and Christine Ranney. 1999. "The Effect of Income and Food Programs on Dietary Quality: A Seemingly Unrelated Regression Analysis with Error Components." *American Journal of Agricultural Economics*, 81: 959-71.
- Wilde, Park, Paul McNamara, and Christine Ranney. 2000. *The Effect on Dietary Quality of Participation in the Food Stamp and WIC Programs*. Washington DC: US Department of Agriculture, ERS Food Assistance and Nutrition Research Report No. 9, September.
- Williams, Richard. 2011. Class notes, Categorical Data Analysis. Fall Semester. University of Notre Dame. <<http://www3.nd.edu/~rwilliam/stats3/>>
- Wood, Ben, Carl Nelson and Lia Nogueira. "Food Price Crisis: Welfare Impact on Mexican Households." Paper presented at the *International Agricultural Trade Research Consortium Analytical Symposium* June 22-23, 2009. Seattle Washington.
- Wrigley, Neil, Daniel Warm, Barrie Margetts, and Amanda Whelan. 2002. "Assessing the Impact of Improved Retail Access on Diet in a 'Food Desert': A Preliminary Report." *Urban Studies*, 39: 2061-2082.
- You, Zhikang, James Epperson, and Chung Huang (2011). "Consumer Demand for Fresh Fruits and Vegetables in the United States (1960 – 1993)". The University of Georgia Cooperative Extension RB 431.

## List of Figures

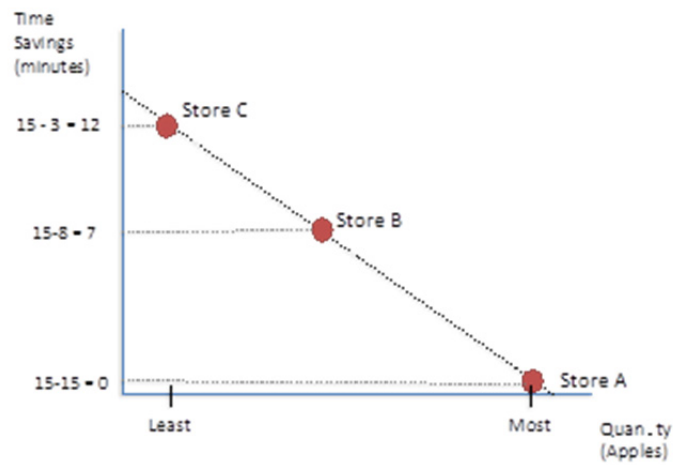
**Figure 1:** Graphical illustration of consumer demand:  
The effect of budget or price ratio change on consumption



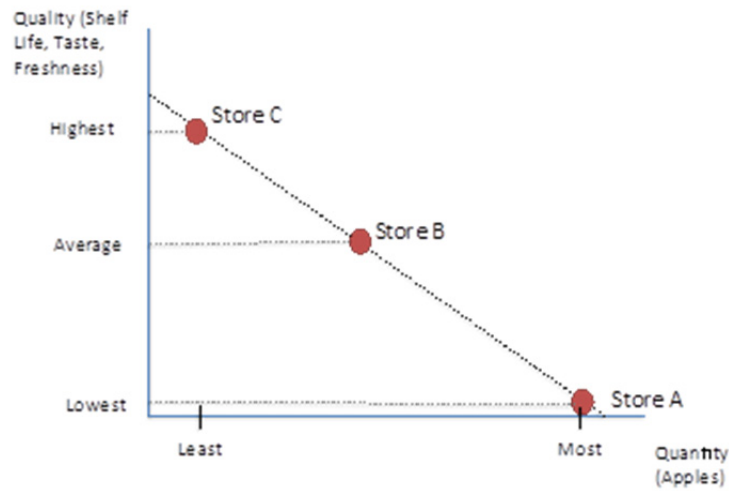
**Figure 2:** Linear and quadratic Engel Curves of food budget shares on log expenditure



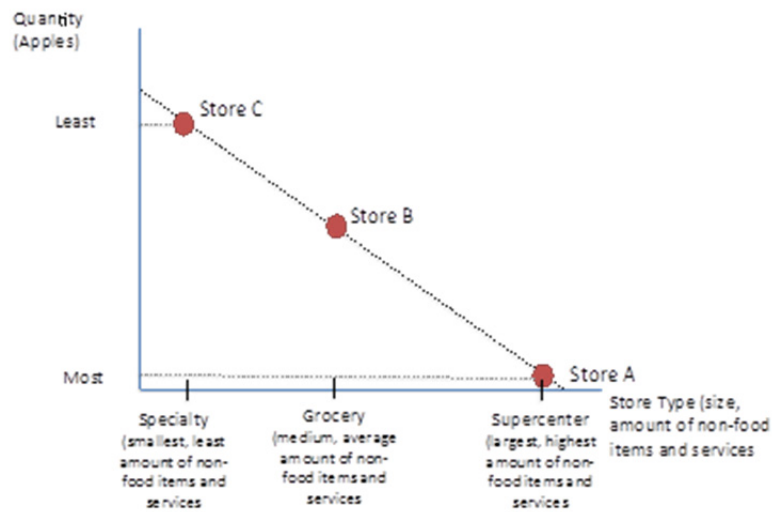
**Figure 3: Tradeoff of time savings vs. quantity**



**Figure 4: Tradeoff of quality vs. quantity**




**Figure 5: Tradeoff of quantity vs. store type**



**Figure 6:** Choice Experiment attributes and levels

Attributes	Levels	Description	Examples
Store Type	Specialty Store	Small-sized store with few non-food items and no additional services (pharmacy, photo, florist, etc.)	<ul style="list-style-type: none"> <li>• Strawberry Fields</li> <li>• World Harvest</li> <li>• corner store</li> </ul>
	Grocery Store	Medium-sized store with an average amount of non-food items and few additional services	<ul style="list-style-type: none"> <li>• Schnucks</li> <li>• County Market</li> </ul>
	Supercenter	Large-sized store with many non-food items and many additional services	<ul style="list-style-type: none"> <li>• Sam's Club</li> <li>• Wal-Mart</li> <li>• Meijer</li> </ul>
Produce Quality & Variety	Limited	Limited Fresh Produce Limited canned or frozen fruits & vegetables	
	Average	Average quality fresh produce Some canned or frozen fruits & vegetables	
	High	High quality fresh produce Variety of canned or frozen fruits & vegetables	
Weekly Grocery Bill	85% 90% 95% 105% 110% 115%	Total grocery bill spent at the store for one week's worth of groceries presented in terms of percentage of respondent's current average weekly grocery bill	
Travel time	5 minutes 15 minutes	Total travel time required to get to the store using the mode of transportation most often used by the respondent	

Figure 7: Sample Choice Experiment question



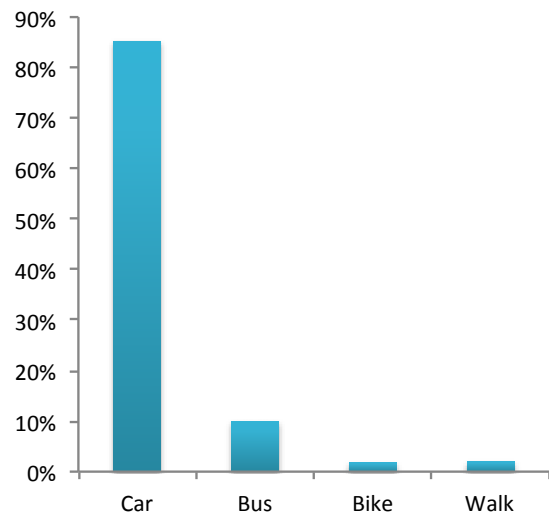
Q2.3.

		Store A		Store B	Current Store
Travel Time		5 minutes from home		15 minutes from home	I would choose my current store.
Weekly Grocery Bill		\$ 59.5		\$ 73.5	
Store Type		Grocery Store		Super Center	
Produce		Limited fresh produce; Limited canned & frozen, fruits & vegetables.		High quality fresh produce; Variety of canned, & frozen fruits & vegetables	

Q2.4. I would choose...

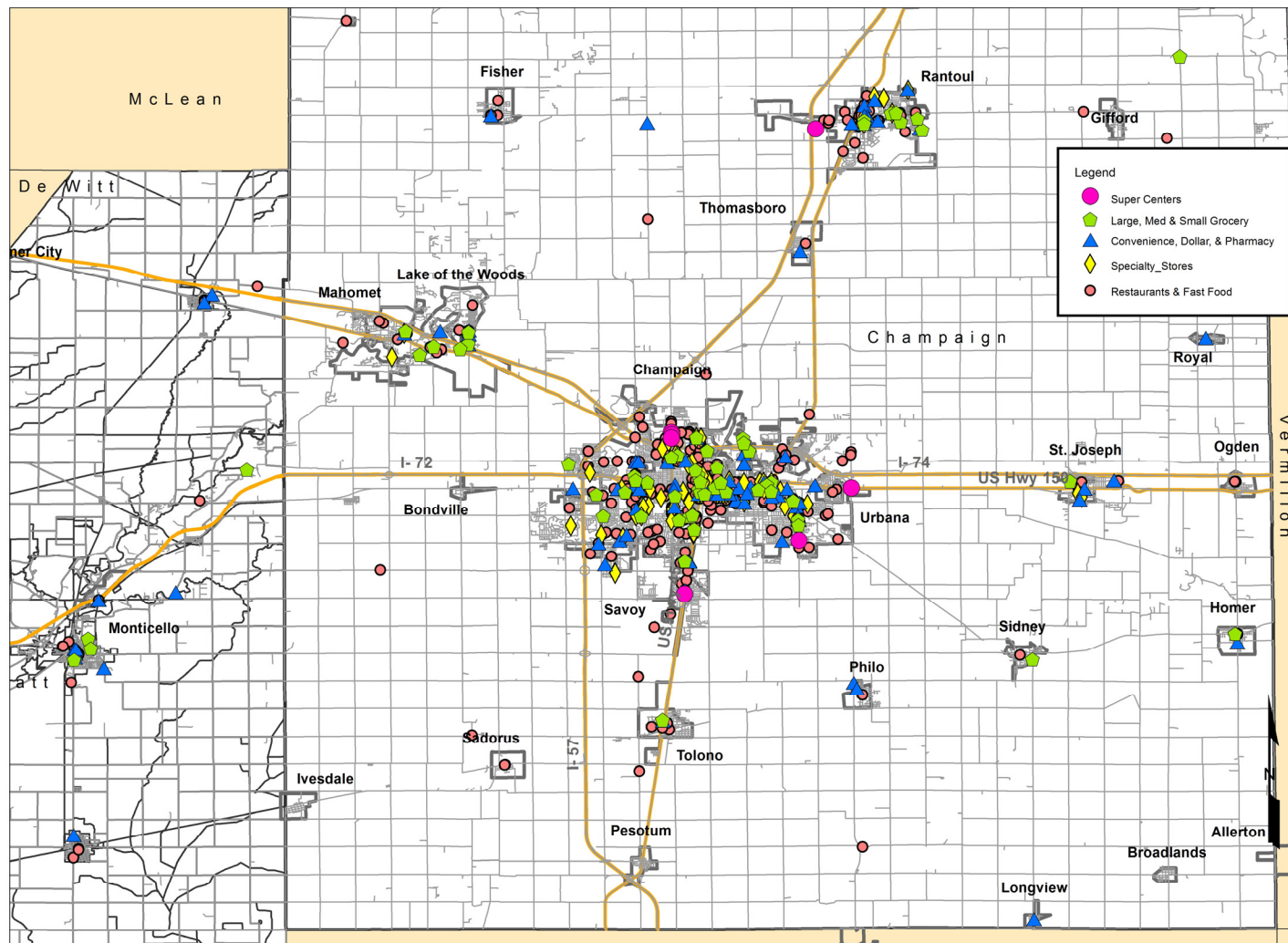
☐ Store A  
☐ Store B  
☐ Current Store

**Figure 8:** Modes of transport most often used by respondents to get to their main grocery store



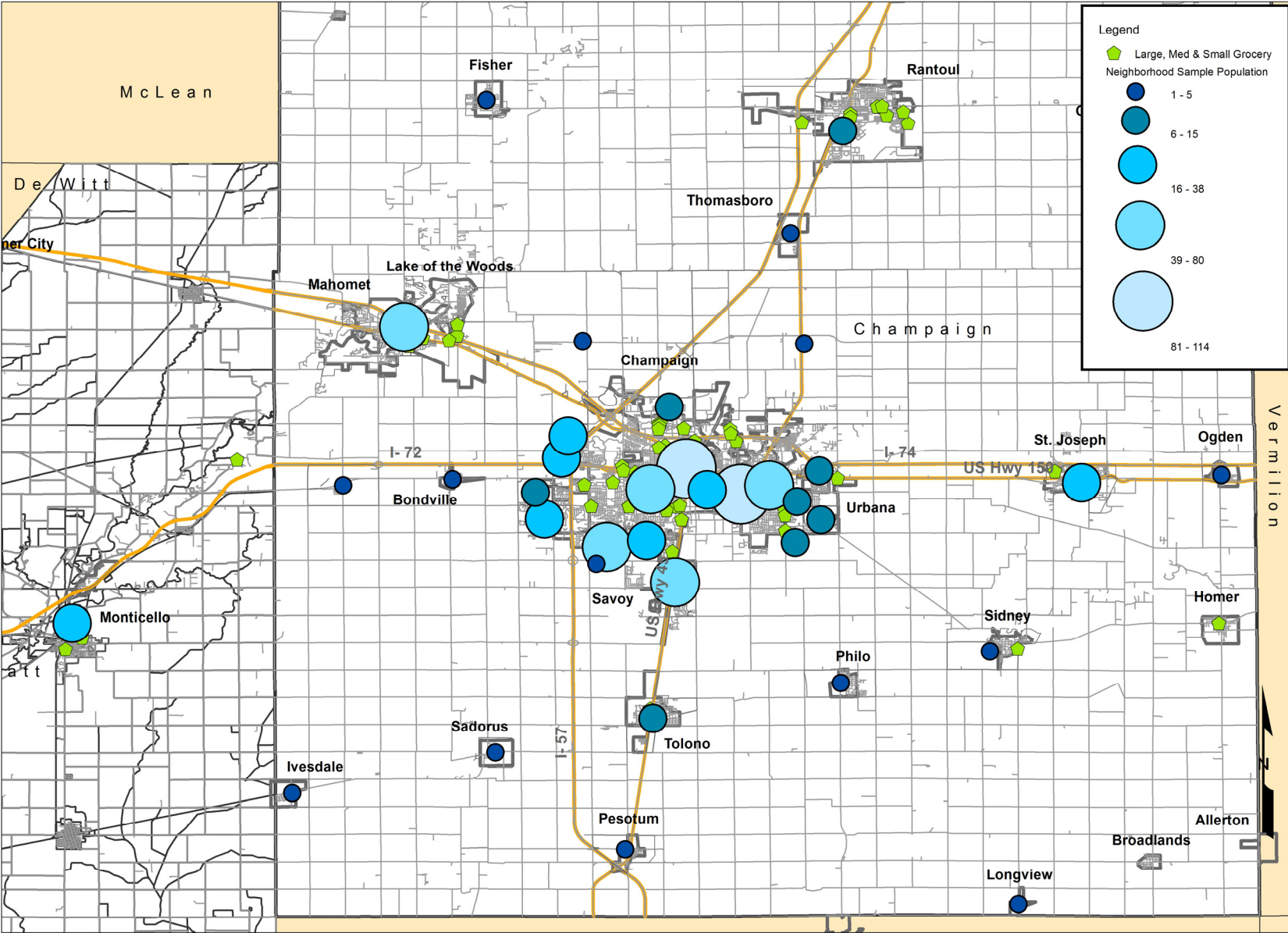


**Figure 9: Food geography of Champaign & Piatt counties**



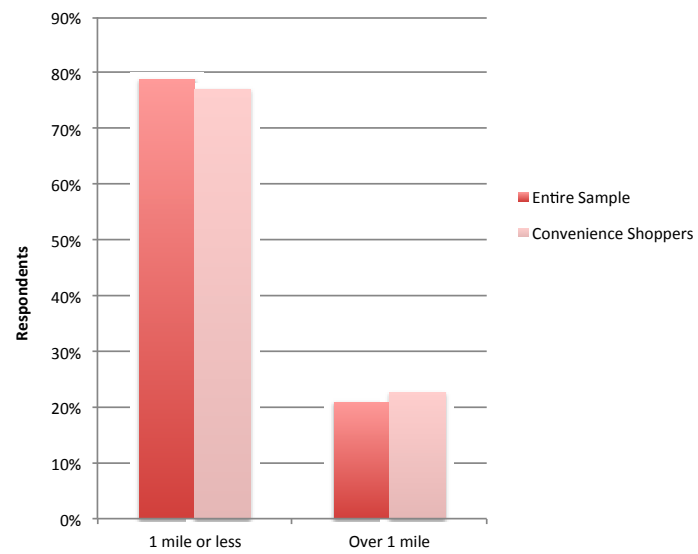
Source: American FactFinder, Census 2012 TIGER, LexisNexis

Figure 10: Sample residential distribution and grocery store locations in Champaign & Piatt counties

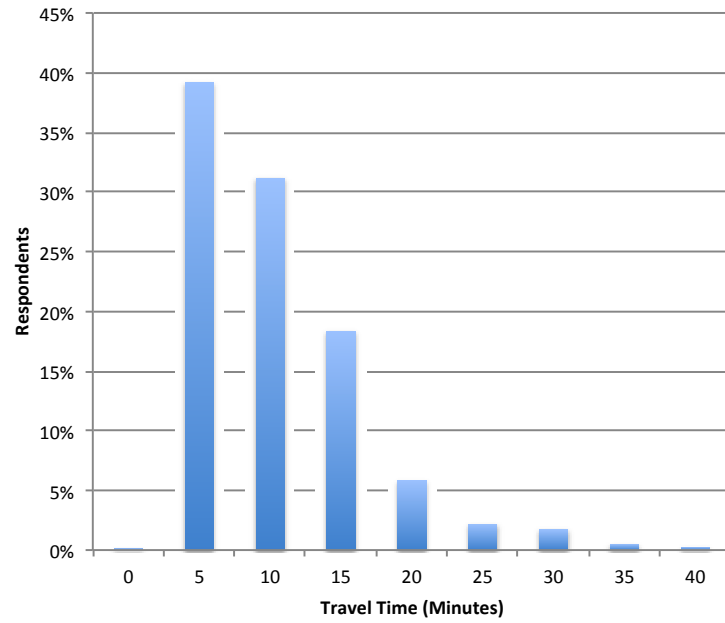


Source: American FactFinder, Census 2012 TIGER, LexisNexis

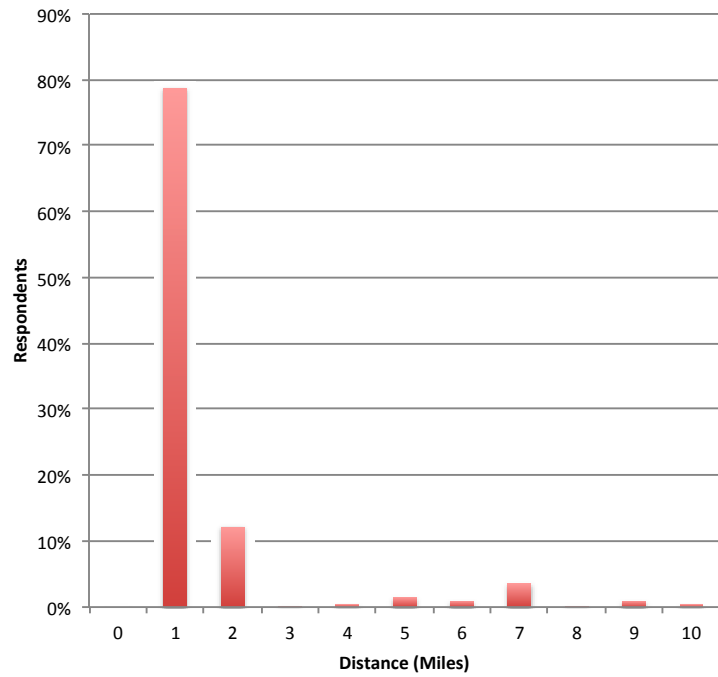
**Figure 11:** Distribution of geographic distance to nearest grocery store:  
comparison of convenience store shoppers



**Figure 12:** Distribution of stated travel time to principal grocery stores



**Figure 13:** Distribution of geographic distance to nearest grocery store



# **Appendix A:** Choice experiment design SAS macro and SAS fractional factorial design output

```
%mktruns (2 3 3 6)
%mktext (2 3 3 6, n=36)
%mktlab (data=design, int=f1-f2)
proc print; run;
%choicetf (data=final, model=class(x1-x4), nsets=18, maxiter=20,
flags=f1-f2, beta=zero);
proc print; by set; id set; run;
```

Profile	Set	Attribute Code				Attribute Description			
		X1	X2	X3	X4	Time	Store	Produce	Price
1	1	1	2	3	1	5 min	Grocery	Limited	- 15%
2	1	2	3	1	4	15 min	SuperCenter	HQ Fresh	+5%
3	2	1	3	2	1	5 min	SuperCenter	Avg Fresh	- 15%
4	2	2	2	1	6	15 min	Grocery	HQ Fresh	+15%
5	3	2	2	3	2	15 min	Grocery	Limited	-10%
6	3	1	3	1	3	5 min	SuperCenter	HQ Fresh	-5%
7	4	1	2	2	3	5 min	Grocery	Avg Fresh	-5%
8	4	2	3	3	5	15 min	SuperCenter	Limited	+10%
9	5	1	1	3	4	5 min	Specialty	Limited	+5%
10	5	2	3	2	1	15 min	SuperCenter	Avg Fresh	- 15%
11	6	2	3	1	3	15 min	SuperCenter	HQ Fresh	-5%
12	6	1	1	2	6	5 min	Specialty	Avg Fresh	+15%
13	7	1	2	3	2	5 min	Grocery	Limited	-10%
14	7	2	1	2	6	15 min	Specialty	Avg Fresh	+15%
15	8	1	2	2	4	5 min	Grocery	Avg Fresh	+5%
16	8	2	1	3	3	15 min	Specialty	Limited	-5%
17	9	2	3	2	2	15 min	SuperCenter	Avg Fresh	-10%
18	9	1	2	1	5	5 min	Grocery	HQ Fresh	+10%
19	10	1	3	3	6	5 min	SuperCenter	Limited	+15%
20	10	2	1	1	2	15 min	Specialty	HQ Fresh	-10%
21	11	1	3	1	4	5 min	SuperCenter	HQ Fresh	+5%
22	11	2	1	2	5	15 min	Specialty	Avg Fresh	+10%
23	12	2	2	1	5	15 min	Grocery	HQ Fresh	+10%
24	12	1	1	3	3	5 min	Specialty	Limited	-5%
25	13	1	1	1	1	5 min	Specialty	HQ Fresh	- 15%
26	13	2	2	2	3	15 min	Grocery	Avg Fresh	-5%
27	14	2	1	1	1	15 min	Specialty	HQ Fresh	- 15%
28	14	1	3	3	5	5 min	SuperCenter	Limited	+10%
29	15	1	1	2	5	5 min	Specialty	Avg Fresh	+10%
30	15	2	3	3	6	15 min	SuperCenter	Limited	+15%
31	16	2	1	3	4	15 min	Specialty	Limited	+5%
32	16	1	2	1	6	5 min	Grocery	HQ Fresh	+15%
33	17	1	3	2	2	5 min	SuperCenter	Avg Fresh	-10%
34	17	2	2	3	1	15 min	Grocery	Limited	- 15%
35	18	2	2	2	4	15 min	Grocery	Avg Fresh	+5%
36	18	1	1	1	2	5 min	Specialty	HQ Fresh	-10%

## Appendix B: Conditional Logit regression results

CLOGIT		
Variable	Coef (SE)	
Main Effects		
ASC	-1.769	***
	0.073	
Weekly Grocery Bill	-0.018	***
	0.004	
Travel Time	-0.073	***
	0.005	
Specialty Store	0.515	***
	0.074	
Grocery Store	0.078	
	0.059	
HQ Produce	1.680	***
	0.292	
AQ Produce	-0.346	
	0.274	
Interaction Effects		
Distance x Grocery Bill	-0.002	**
	0.001	
CurrentTravel x Grocery	0.000	
	0.000	
Income x HQ Produce	0.164	***
	0.054	
Income x AQ Produce	0.176	***
	0.050	
Observations		
	9,609	
Log Likelihood	-2492.06	
Prob > Chi2	0.000	
AIC	5006.13	
BIC	5085.00	
*** significant at 1%; ** significant at 5%,		
* significant at 10%		

**Appendix C:** List of produce for which respondents provided consumption information

<b>Vegetables</b>	<b>Fruits</b>
Potato	Melon
Onion	Berries
Fresh Lettuce	Apples
Leafy Greens	Bananas
Pre-packaged Lettuce	Oranges
Tomato	Tangerines
Cucumber	Lemon
Bell Pepper	Lime
Broccoli	Grapes
Cauliflower	Stone Fruit
Celery	Clementine
Carrots	Pears
Green Beans	Kiwi
Squash	Pineapple
Zucchini	Grapefruit
Asparagus	Pomegranate
Beets	Frozen Fruit
Avocado	Canned Fruit
Snap Peas	Canned Pineapple
Mushrooms	
Sweet Potatoes	
Brussel Sprouts	
Leeks	
Green Onion	
Garlic	
Fennel	
Celeriac	
Turnip	
Parsnip	
Herbs	
Ginger	
Olives	
Canned Vegetables	
Pasta Sauce	
Frozen Vegetables	

**Appendix D:** Distribution of choice experiment selections based on the produce quality attribute

<b>Constrained Group</b>		
Produce Quality	Status Quo (C')	Atlernative (A or B)
Low or Avg	2.4%	28.4%
High	14.9%	54.3%

<b>Full Sample</b>		
Produce Quality	Status Quo (C')	Atlernative (A or B)
Low or Avg	32.0%	23.2%
High	22.0%	22.9%

<b>Unconstrained Group</b>		
Produce Quality	Status Quo (C')	Atlernative (A or B)
Low or Avg	49.2%	2.5%
High	34.5%	13.8%



## Appendix E: MMNL results with Main Effects - Only specification

Econometric specification of main effects-only model:

$$U_{ij} = \alpha_j + \beta_1 X_{iBill} + \beta_2 X_{iTravel} + \beta_3 X_{iSpecialty} + \beta_4 X_{iGrocery} + \beta_5 X_{iHQProduce} + \beta_6 X_{iAQProduce} + \varepsilon_{ij}$$

### Case 1: Geographic Constraint Classification

Variable	MMNL Main Effect - Full Sample			MMNL Main Effect - Geographic Classification		
	Coef (SE)		SD Pval <sup>a</sup>	Coef (SE)		SD Pval <sup>a</sup>
<b>Main Effects</b>						
ASC	-2.096 ***			-2.186 ***		
	0.074			0.106		
Weekly Grocery Bill	-0.040 ***			-0.038 ***		
	0.003			0.004		
Travel Time	-0.123 ***	0.000		-0.134 ***	0.000	
	0.007			0.011		
Specialty Store	0.428 ***	0.000		0.400 ***	0.000	
	0.101			0.145		
Grocery Store	0.045	0.000		-0.011	0.000	
	0.078			0.111		
HQ Produce	3.281 ***	0.000		3.294 ***	0.000	
	0.124			0.173		
AQ Produce	0.912 ***	0.000		0.800 ***	0.000	
	0.112			0.156		
<b>Main Effect 'Food Desert' Dummies</b>						
ASC_FD				-0.309	0.000	
				0.323		
Weekly Grocery Bill_FD				-0.027 **	0.000	
				0.012		
Travel Time_FD				-0.004	0.197	
				0.023		
Specialty Store_FD				0.291	0.108	
				0.360		
Grocery Store_FD				0.172	0.073	
				0.275		
HQ Produce_FD				1.171 ***	0.000	
				0.418		
AQ Produce_FD				0.135	0.455	
				0.335		
<b>Observations</b>						
	15,534			10,098		
Log Likelihood	-3728.66			-2385.23		
Prob > Chi2	0.000			0.000		
LR Chi 2 (9)	598.39			441.22		
AIC	7471.33			4798.47		
BIC	7524.88			4899.55		

\*\*\* significant at 1%; \*\* significant at 5%, \* significant at 10%

<sup>a</sup>Significance of standard deviations of individual heterogeneity

## Case 2: Preference-based constraint classification

Main Effects MMNL - Preference Classification			
Variable	Coef (SE)		SD Pval <sup>a</sup>
<b>Main Effects</b>			
ASC	-4.155 *** 0.272		
Weekly Grocery Bill	-0.032 *** 0.006		
Travel Time	-0.132 *** 0.019		0.001
Specialty Store	0.174 0.225		0.000
Grocery Store	-0.219 0.162		0.975
HQ Produce	3.151 *** 0.329		0.008
AQ Produce	0.264 0.384		0.687
<b>Main Effect 'Constrained' Dummies</b>			
ASC_CD	3.633 *** 0.378		0.189
Weekly Grocery Bill_CD	-0.041 *** 0.016		0.001
Travel Time_CD	0.052 ** 0.027		0.475
Specialty Store_CD	0.363 0.440		0.287
Grocery Store_CD	0.308 0.318		0.258
HQ Produce_CD	1.292 ** 0.662		0.061
AQ Produce_CD	0.573 0.506		0.277
<b>Observations</b>			
Log Likelihood	4,866		
Prob > Chi2	-734.50		
AIC	0.004		
BIC	1497.01		
	1587.87		

\*\*\* significant at 1%; \*\* significant at 5%, \* significant at 10%

<sup>a</sup>Significance of standard deviations of individual heterogeneity

Case 3: “Constrained” vs. “Opt to Travel” comparison

MMNL Main Effects - Opt to Travel			
Variable	Coef (SE)		SD Pval <sup>a</sup>
Main Effects			
ASC	-0.338 0.240		
Weekly Grocery Bill	-0.059 0.010	***	
Travel Time	-0.077 0.019	***	0.117
Specialty Store	0.426 0.345		0.000
Grocery Store	0.271 0.271		0.000
HQ Produce	4.137 0.459	***	0.000
AQ Produce	0.873 0.294	***	0.185
Main Effect 'Opt to Travel' Dummies			
ASC_OCD	-2.320 0.313	***	0.000
Weekly Grocery Bill_OCD	0.009 0.012		0.000
Travel Time_OCD	-0.049 0.024	**	0.016
Specialty Store_OCD	-0.241 0.391		0.037
Grocery Store_OCD	-0.242 0.311		0.074
HQ Produce_OCD	-0.486 0.468		0.138
AQ Produce_OCD	-0.308 0.346		0.192
Observations			
Log Likelihood	5,601		
Prob > Chi2	-1323.36		
LR Chi 2 (8)	0.000		
AIC	260.93		
BIC	2674.73		
	2767.56		

\*\*\* significant at 1%; \*\* significant at 5%, \* significant at 10%

<sup>a</sup>Significance of standard deviations of individual heterogeneity

**Appendix F: Weekly Grocery Bill estimators linear combination significance test by respondent group**

Econometric specification (equation 12):

$$U_{ij} = \alpha_j + \beta_1 X_{iBill} + \beta_2 X_{iTravel} + \beta_3 X_{iSpecialty} + \beta_4 X_{iGrocery} + \beta_5 X_{iHQProduce} + \beta_6 X_{iAQProduce} + \beta_7 X_{iBill} * Y_{iCurrentTravel} + \beta_8 X_{iBill} * Y_{iDistance} + \beta_9 X_{iHQProduce} * Y_{iIncome} + \beta_{10} X_{iAQProduce} * Y_{iIncome} + \varepsilon_{ij}$$

	"1 mile or less"	"Over 1 mile"	"Unconstrained"	"Constrained"	"Opt to Travel"
Equation ( $dU/dBill$ )	$X_{Bill} + (X_{Bill} * Y_{Distance}) + (X_{Bill} * Y_{CurrentTravel}) = 0$	$X_{Bill} + X_{Bill\_FD} + (X_{Bill} * Y_{Distance}) + (X_{Bill} * Y_{CurrentTravel}) = 0$	$X_{Bill} + (X_{Bill} * Y_{Distance}) + (X_{Bill} * Y_{CurrentTravel}) = 0$	$Y_{Distance} + (X_{Bill} * Y_{CurrentTravel}) = 0$	$X_{Bill} + X_{Bill\_OCD} + (X_{Bill} * Y_{Distance}) + (X_{Bill} * Y_{CurrentTravel}) = 0$
Calculation	$-0.031 + (-0.01 * 0.5) + (-0.001 * 8.6) = 0$	$-0.031 + 0.003 + (-0.01 * 4.3) + (-0.001 * 12.3) = 0$	$-0.027 + (-0.010 * 0.5) + (0.000 * 7.4)$	$-0.027 + -0.014 + (-0.010 * 4.1) + (0.000 * 12.3)$	$-0.033 + -0.008 + (-0.009 * 0.4) + (-0.001 * 12.1)$
<i>'Weekly Grocery Bill' Point Estimate Results</i>					
Coefficient	-0.0451***	-0.082***	-0.032***	-0.082***	-0.081***
Standard Error	0.005	0.013	0.007	0.013	0.015
P-Value	0.000	0.000	0.000	0.000	0.000

## Appendix G: Regression results of Robustness Tests

### Test 1: MMNL Results using 2-mile cut-off

	MMNL w/ dummy 2 mile cut-off		
Variable	Coef	(SE)	SD Pval <sup>a</sup>
<b>Main Effects</b>			
ASC	-2.366	***	
	0.116		
Weekly Grocery Bill	-0.033	***	
	0.008		
Travel Time	-0.139	***	0.000
	0.012		
Specialty Store	0.558	***	0.000
	0.149		
Grocery Store	0.041		0.000
	0.114		
HQ Produce	2.367	***	0.000
	0.459		
AQ Produce	-0.618		0.000
	0.429		
<b>Main Effect 'Food Desert' Dummies</b>			
ASC_FD	-0.028		0.002
	0.788		
Weekly Grocery Bill_FD	-0.043		0.880
	0.048		
Travel Time_FD	-0.024		0.000
	0.051		
Specialty Store_FD	-1.062	*	0.002
	0.612		
Grocery Store_FD	-0.370		0.117
	0.452		
HQ Produce_FD	2.116	***	0.001
	0.789		
AQ Produce_FD	0.794		0.606
	0.563		
<b>Interaction Effects</b>			
Distance x Grocery Bill	-0.002		0.379
	0.007		
CurrentTravel x Grocery Bill	-0.001	***	0.000
	0.001		
Income x HQ Produce	0.256	***	0.556
	0.087		
Income x AQ Produce	0.278	***	0.064
	0.081		
<b>Observations</b>			
Log Likelihood	9,609		
Prob > Chi2	-2246.61		
	0.000		
AIC	4529.21		
BIC	4658.28		

\*\*\* significant at 1%; \*\* significant at 5%, \* significant at 10%

<sup>a</sup>Significance of standard deviations of individual heterogeneity

## Test 2: MMNL Results with lower-income respondent sub-sample

	MMNL w/ dummy Lower Income Sample		
Variable	Coef	(SE)	SD Pval <sup>a</sup>
<b>Main Effects</b>			
ASC	-2.108	***	
	0.188		
Weekly Grocery Bill	-0.088	***	
	0.018		
Travel Time	-0.141	***	0.000
	0.019		
Specialty Store	0.422	***	0.007
	0.241		
Grocery Store	-0.319		0.000
	0.209		
HQ Produce	3.314	***	0.034
	1.177		
AQ Produce	-1.684		0.000
	1.087		
<b>Main Effect 'Food Desert' Dummies</b>			
ASC_FD	-5.509	**	0.085
	2.164		
Weekly Grocery Bill_FD	-0.155	**	0.536
	0.076		
Travel Time_FD	-0.049		0.001
	0.066		
Specialty Store_FD	1.020		0.000
	1.082		
Grocery Store_FD	0.855		0.002
	0.868		
HQ Produce_FD	2.399		0.000
	1.497		
AQ Produce_FD	-0.051		0.002
	1.120		
<b>Interaction Effects</b>			
Distance x Grocery Bill	0.016		0.105
	0.013		
CurrentTravel x Grocery Bill	0.001		0.000
	0.002		
Income x HQ Produce	0.096		0.005
	0.338		
Income x AQ Produce	0.748	**	0.760
	0.315		
<b>Observations</b>			
Log Likelihood	3,207		
Prob > Chi2	-760.17		
	0.000		
AIC	1556.35		
BIC	1665.67		

\*\*\* significant at 1%; \*\* significant at 5%, \* significant at 10%

<sup>a</sup>Significance of standard deviations of individual heterogeneity

Test 3: MMNL Results with broader "Constrained" respondent group

Variable	MMNL w/ dummy "Constrained" def. 2		
	Coef (SE)		SD Pval <sup>a</sup>
<b>Main Effects</b>			
ASC	-4.206 ***		
	0.288		
Weekly Grocery Bill	-0.025 ***		
	0.010		
Travel Time	-0.133 ***		0.001
	0.019		
Specialty Store	0.028		0.002
	0.262		
Grocery Store	-0.161		0.757
	0.167		
HQ Produce	3.134 ***		0.682
	0.786		
AQ Produce	0.028		0.719
	0.776		
<b>Main Effect 'Constrained' Dummies</b>			
ASC_CD	3.070 ***		0.007
	0.370		
Weekly Grocery Bill_CD	0.006		0.970
	0.014		
Travel Time_CD	0.079 ***		0.124
	0.024		
Specialty Store_CD	0.892 **		0.329
	0.425		
Grocery Store_CD	0.557 **		0.466
	0.275		
HQ Produce_CD	0.494		0.077
	0.568		
AQ Produce_CD	0.469		0.768
	0.488		
<b>Interaction Effects</b>			
Distance x Grocery Bill	-0.008 ***		0.149
	0.003		
CurrentTravel x Grocery Bill	-0.001		0.014
	0.001		
Income x HQ Produce	0.031		0.175
	0.129		
Income x AQ Produce	0.046		0.797
	0.118		
<b>Observations</b>			
	5,058		
Log Likelihood	-840.30		
Prob > Chi2	0.0014		
AIC	1716.61		
BIC	1834.13		

\*\*\* significant at 1%; \*\* significant at 5%, \* significant at 10%

<sup>a</sup>Significance of standard deviations of individual heterogeneity